ISBN: 978-90-819670-0-6
Printed by: Print Service Ede
Proofread by: Robert John Correll

Deciphering the genetic background of Systemic Sclerosis

Proefschrift

Ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. mr. S.C.J.J. Kortmann
volgens het besluit van het college van decanen
in het openbaar te verdedigen op
dinsdag 9 oktober 2012 om 10.30 uur.

door

Jacobus Christiaan Andreas Broen
Geboren op 7 april 1985
te Linne

Promotor:	Prof. dr. P.L.C.M. van Riel
Co-promotoren:	Dr. T.R.D.J. Radstake Dr. M.J.H. Coenen
Manuscriptcommissie:	Prof. dr. M. Netea (voorzitter) Prof. dr. H.G. Brunner Prof. dr. B. Roep (LUMC)

Cover:

Contemplation (1938) Paul Klee (1879–1940).
Artwork is in the public domain.

Klee was an influential 20th century abstract painter. At the age of 57 he was diagnosed with progressive scleroderma. After diagnosis, there was a steep decline in productivity and marked change in style. However, he still managed to produce more than 1000 works in the last years of his life. A number of these are thought to be inspired by his disease and reflect fear, suffering and death. Other drawings reflect his view on scleroderma and the disfigurement of his body.

Although "Contemplation" does not seem to be influenced by his disease, the meanings of the hieroglyphic-like lines in the painting are still as enigmatic as the processes underlying systemic sclerosis. Hopefully this thesis, contemplating genetics, will provide a key to decipher a small part of this disease.

Whether Klee would be happy with this comparison is unlikely, considering his quotation: "The worst state of affairs is when science begins to concern itself with art."

....for Jac Broen

Table of contents

Chapter 1

General Introduction

General introduction

A brief history

Although systemic sclerosis (SSc) is a rare disease, the characteristic feature of hardened skin (scleroderma) has been mentioned throughout the history of medicine. Scleroderma originates from the Greek words skleros (hard or indurated) and derma (skin). Hippocrates (400 B.C.) first used the term scleroderma, for describing hardened skin (Aphorism V:71). Hippocrates notes "In those persons in whom the skin is stretched, parched, and hard, the disease terminates without sweats." Whether Hippocrates really encountered a patient with SSc is not deductible from his notations (1). A more precise description was not provided until 1753. Doctor Carlo Curzio from Naples describes in a published monograph a 17-year-old woman with symptoms of hardness of the skin (differing in degree from place to place), tightness around the mouth, and hardness around the neck. He also noted loss of warmth in the skin, but normal pulse, respiration and digestion. Intriguingly, and in sharp contrast to the current clinical practice, Dr. Curzio was able to cure the disease fully in this female, by prescribing warm milk and vapor baths, bleeding from the foot, and small doses of quicksilver. After 11 months the skin became soft again (2). Although more reports on scleroderma followed, it was not until 1945 that Robert H. Goetz first described the concept of scleroderma as a systemic disease. To accentuate this, he introduced the term progressive systemic sclerosis (3).

Epidemiology

SSc is a rare connective tissue disease and prevalence varies geographically (4). This fact initially hampered thorough epidemiological studies. More extensive research illustrating the prevalence, incidence and risk factors for SSc appeared after the first common disease description for SSc was published by the American Rheumatism Association in 1980 (**Table 1**) (5). Now, prevalence is estimated between 3 and 24 per 100,000 persons and interestingly, appears to be higher in the United States and Australia compared to Japan and Europe (6-10). In Europe there seems to be a North to South gradient with North Europeans having a lower prevalence of SSc (8). One of the most interesting populations is the Choctaws

Indians, a Native American tribe in Oklahoma. An extremely high prevalence of SSc was obsered (469/100,000) based on 14 SSc cases over a four year period in full blood Choctaws. This prevalence was significantly higher than the prevalence observed in non Choctaws Native Americans in Oklahoma (9.5/100,000) (11). Noteworthy, no significant contributing environmental factor was found in a case–control study and a role for a genetic component was confirmed by the association of the HLA region of these Native Americans with disease risk. SSc is present more frequently in women, with ratios varying from 3:1 to 14:1 (8). This finding may be explained by differences in environmental exposures, hormonal milieu and the different genetic makeup in the sense of harboring two X chromosomes (12, 13). In addition to the higher incidence rate of disease in females, SSc is more frequently found in populations originating from Africa. For instance, a study conducted in the USA found a SSc incidence of 3.1/100,000/year in black women in disparity to 2.7/100.000/year in white women (14). Interestingly, the diffuse cutaneous systemic sclerosis (dcSSc) subform accounted for 60% versus 27% of the cases in black compared to white SSc females (7). Furthermore, black patients developed SSc at an earlier age than Caucasian patients. A study performed in France postulated similar findings (15).

Table 1. ARA/ACR criteria for SSc

Criteria	Phenotype
Major:	• Proximal scleroderma
Minor (2):	• Sclerodactyly • Digital pitting scars/loss of substance of finger pads • Bibasilar pulmonary fibrosis

Pathogenesis and pathophysiology

Systemic sclerosis is an autoimmune disease characterized by extensive fibrosis of the skin and internal organs. These hallmark manifestations comprise vasculopathy, immune system activation and exaggerated collagen deposition in extracellular matrix (fibrosis). Although the disease has an impressive appearance

and high mortality, the pathogenesis remains poorly understood. Endothelial cell activation, vascular damage and immune activation are generally regarded as the first events. The vascular damage is followed by the extravasation of inflammatory cells, including monocytes and lymphocytes. Eventually, fibroblasts are activated. These activated fibroblasts start producing excessive extracellular matrix, commencing the fibrotic features of SSc. In time, when looking in advanced lesional skin, inflammation seems to be halted again (16-22). Environmental, genetic and epigenetic factors are thought to underlie the onset of SSc and are discussed in the following paragraphs. A schematic view of the processes involved is provided in **Figure 1**.

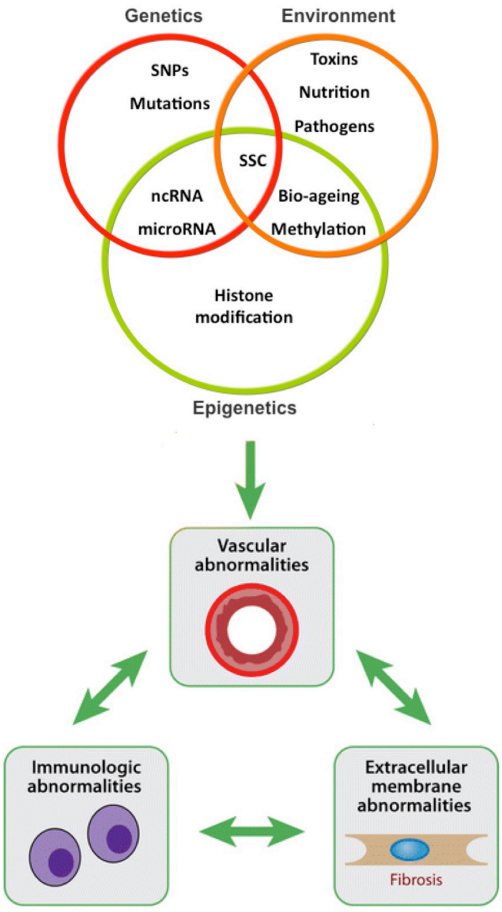


Figure 1. Processes involved in systemic sclerosis pathogenesis

Vascular abnormalities

Vasculopathy in SSc comprises Raynaud's phenomenon, increased vascular wall thickness, vascular occlusion, devascularization, and thickening of the basement membrane. More recently, it was discovered that changes in the nailfold capillaries are one of the first signs in SSc (16,17) Vasculopathy is also the most common process involved in the major complications of SSc, namely myocardial dysfunction, pulmonary arterial hypertension and scleroderma renal crisis (17). Thinning of the media, intimal proliferation and fibrosis of the adventitia strike arterioles during the course of the disease. Next to these processes there is reduced vasodilatory capacity, hypoxia, oxidative stress and increased adhesiveness of the blood vessels to lymphocytes (16-18). Endothelial cells play an undefined role in SSc pathogenesis; they seem to be both targets of immune activity and immune activity mediators (18).

The immune cell infiltrates around blood vessels of SSc patients mainly consist of CD4+ Tcells, consequently these cells are the best studied immune cells in SSc (19,20). Of interest, T regulatory cells, the immune activity-diminishing counterparts of CD4+ cells, seem to be impaired in SSc (21). CD4+ T cell cytokines are, in concert with monocyte-derived cytokines, elevated in the serum of SSc patients. These cytokines include IL-1, IL-2, IL-4 and IL-6. Intriguingly, IL-2 can induce the secretion of active TGFβ by monocytes. TGFβ is able to activate fibroblasts to secrete more extracellular matrix. Activated fibroblasts start producing both cytokines and growth factors including CTGF and TGFβ, CTGF mRNA remains overexpressed in SSc fibroblasts and fibrotic skin lesions (22). This phenomenon might contribute to an autocrine loop; in turn there is increased ICAM-1 expression on fibroblasts, which increases the adhesion of lymphocytes. Although all these processes need further scientific exploration and validation, they currently provide the best explanation for the disease perpetuating processes in SSc taking into account cell-cell interactions and cytokines (16-22). Next to the abovementioned cytokines, IL-8, IL-10, IL-13, IL-17, TGFβ, PDGF, TNF-α, IFN-γ, and IL-2 receptor are increased in the serum of SSs patients as well, but their role in pathogenesis is still a matter of intense scientific debate.

Clinical manifestations

Systemic sclerosis can be divided into two clinical phenotypes, limited and diffuse cutaneous SSc. Both have their own clinical characteristics as comprehensively displayed in **Table 2**. These differences have been used by LeRoy to postulate classification criteria for SSc phenotypes. Briefly, patients with lcSSc have a disease course progressing more slowly and the skin changes do not extend beyond the elbows and knees into the proximal extremities or trunk. Internal organ involvement is highly variable in lcSSc. Patients with dcSSc develop more rapidly progressing skin fibrosis and severe internal organ involvement (23). Autoantibody production is different between subtypes as well; anticentromeric antibodies are seen almost exclusively in 70-80% of the lcSSc patients and anti-topoisomerase antibodies in 30% of dcSSc patients.

Table 2. LeRoy classification of systemic sclerosis phenotypes

SSc subset	Symptom
Diffuse	• Onset of Raynaud's within 1 year of onset of skin changes (puffy or hidebound skin) • Truncal and acral skin involvement • Presence of tendon friction rubs • Early and significant incidence of interstitial lung disease, oliguric renal failure, diffuse gastrointestinal disease, and myocardial involvement • Absence of anticentromere antibodies • Nailfold capillary dilation and capillary destruction • Antitopoisomerase antibodies (30% of patients)
Limited	• Raynaud's phenomenon for years (occasionally decades) • Skin involvement limited to hands, face, feet, and forearms (acral) or absent • A significant late incidence of pulmonary hypertension, with or without interstitial lung disease, trigeminal neuralgia, skin calcifications, telangiectasias • A high incidence of anticentromere antibodies (70%-80%) • Dilated nailfold capillary loops, usually without capillary dropout

Raynaud's phenomenon is a triphasic (pallor, cyanosis and red engorgement) color change of the fingers and toes, which is usually induced by cold or stress.

Raynaud's phenomenon is not always accompanied by SSc development; however, SSc is always preceded by it. In lcSSc, it is present multiple years before skin changes or internal organ complications occur. In dcSSc, it precedes the onset of skin and organ manifestations not more than one year. The first skin changes in the early phase consist of skin edema, mainly affecting fingers and hands. Next to this, painful ischemic digital ulceration frequently develops. Particularly in dcSSc, the edema is rapidly followed by development of firm, indurated skin over the extremities, trunk, and face. In lcSSc the edema and skin thickening are confined to the distal extremities. The skin keeps on thickening until approximately the first 5 years of disease are passed (16,17).

Organ involvement in SSc is widespread. Involvement of lungs, heart and kidney are the most devastating and most deadly complications. For that reason they will be described below. An extensive overview of all complications occurring in SSc is provided in **table 3** (16,17,24).

Table 3. Complications and symptoms of Systemic Sclerosis

System	Manifestation
Cutaneous	• Diffuse edema of hands and feet (early stages) • Progressive skin tightening • Sclerodactyly • Calcinosis • Telangiectasias • Digital ulcers and pits • Contractures • Hyperpigmentation, hypopigmentation, salt and pepper skin • Characteristic facies
Vascular	• Raynaud's phenomenon • Nailfold capillary changes • Digital ischemia and ulcers • Vasculitic leg ulcers
Pulmonary	• Interstitial lung disease, including alveolitis and interstitial fibrosis • Pulmonary hypertension • Recurrent aspiration pneumonitis caused by esophageal reflux and dysmotility • Decreased thoracic compliance • Respiratory muscle weakness

System	Manifestation
Cardiac	• Cardiomyopathy (systolic and diastolic dysfunction): Congestive heart failure • Conduction defects • Septal infarction pattern • Ventricular conduction abnormalities • Arrhythmias • Heart blocks • Pericarditis or pericardial effusion
Renal	• Scleroderma renal crisis (hypertension, renal failure MAHA)
Musculoskeletal and Rheumatologic	• Arthralgia • Tendon friction rubs • Inflammatory arthritis, erosive arthropathy • Myopathy, myositis
Gastrointestinal	• Gastroesophageal reflux • Esophageal dysmotility, aperistaltic esophagus • Esophageal stricture • Watermelon stomach • Decreased peristalsis throughout the GI tract • Bacterial overgrowth and malabsorptive diarrhea, alternating diarrhea and constipation • Megacolon • Colonic wide-mouth diverticuli • Pneumatosis cystoides intestinales • Primary biliary cirrhosis • Anal incontinence
Endocrine	• Hypothyroidism
Neurologic	• Carpal tunnel syndrome • Trigeminal neuralgia

Pulmonary complications may present as dyspnea and nonproductive cough. Especially shortness of breath, fatigue, chest pain, and occasionally syncope are warning signs of pulmonary arterial hypertension, the most feared complication of SSc. Pulmonary arterial hypertension develops in about 20% to 30% of lcSSc patients. Patients with dcSSc, most often develop interstitial lung disease, the pulmonary manifestations might be more severe in dcSSc patients who are ATA positive or have rapidly progressing skin disease. (17, 23) Renal involvement, or scleroderma renal crisis, is characterized by the development of severe

hypertension, renal failure, and microangiopathic hemolytic anemia. The development of renal crisis only occurs in patients with dcSSc (16,17). Cardiac complications are rare, but myocardial fibrosis may develop in patients with dcSSc (24).

Depending on the development of complications, the survival rate of SSc patients drops. Patients with lcSSc have a 10 year survival of 75-79%, only 53%-62% of patients with dcSSc survive the first 10 years after diagnosis. However, when a patient with lcSSc develops PAH, 5 year survival drops to 10% (25-28).

Environmental factors

Exposure to some chemicals seems to result in symptoms that mimic SSc such as polyvinyl chloride intoxication or the toxic oil syndrome (29). An Australian study found that SSc patients more frequently had an occupation involving retail, wholesale, construction, manufacturing and rural activities compared to healthy persons (6). These jobs place an individual at risk for exposure to hydrocarbons and silica dust. A recent meta-analysis addressing the effect of silica exposure on SSc development by analyzing all reports up to 2007 was severely hampered by significant heterogeneity among the studies included (30). Although multiple substances have been investigated, up till now none of them was implicated in SSc susceptibility in a convincing and reproducible matter.

Genetics and epigenetics in Systemic Sclerosis

The observations of environment independent ethnic clustering, the strong predominance of disease in females and the lack of a clear general environmental factor involved in SSc pathogenesis have inspired several researchers to decipher the genetic background for SSc susceptibility. Genetic research has been performed on the level of heritability studies in families, chromosomal, genome wide and candidate gene studies. The next paragraphs discuss the findings coming forth from these efforts. An overview is provided in **figure 2**.

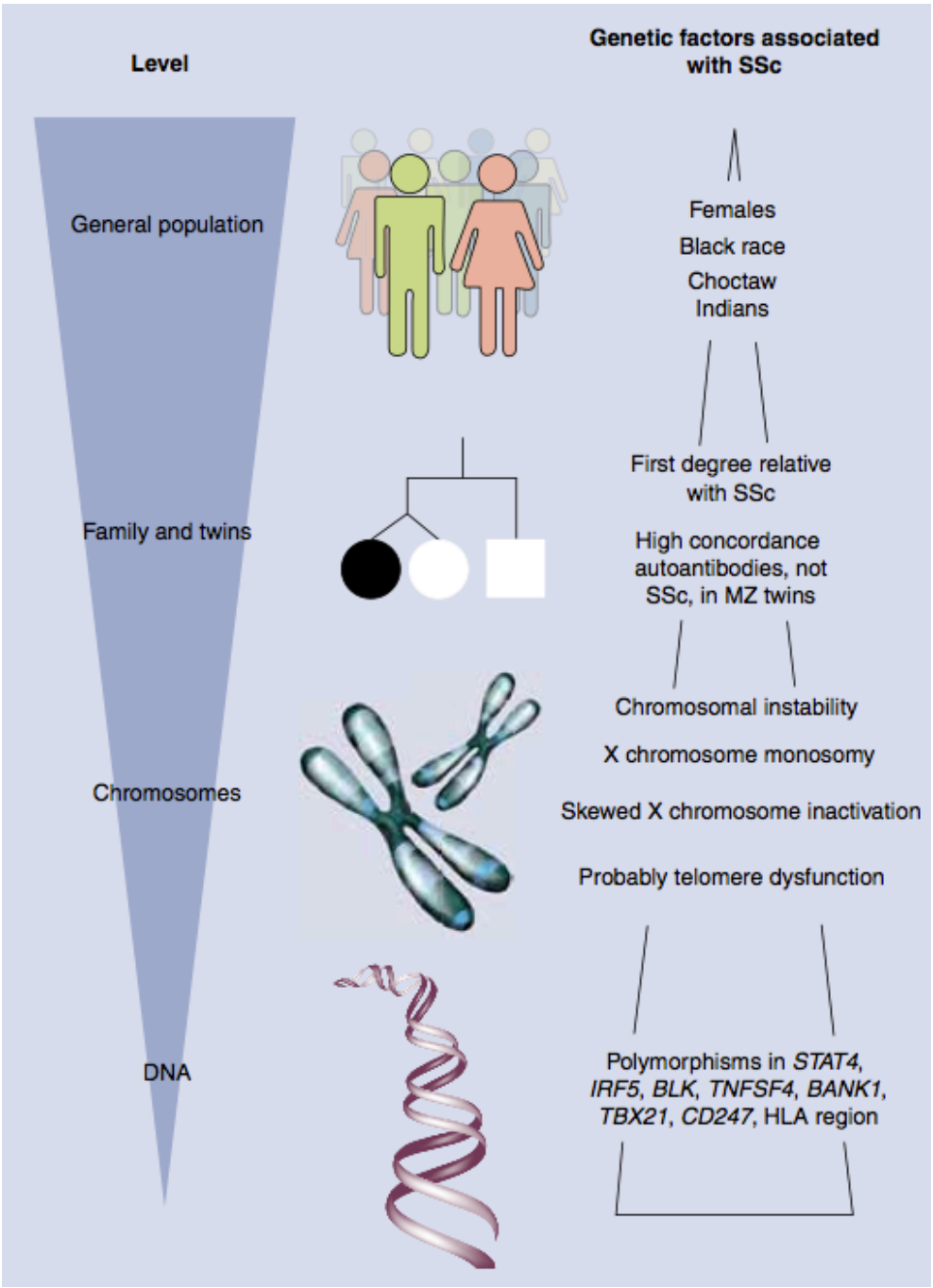


Figure 2. Multilevel overview of genetic involvement in SSc susceptibility (116)

Family studies

Family studies are the first logical step to investigate heritability. These studies have shown that indeed genetic factors play a role in SSc pathogenesis. Two studies investigated whether having a positive family history contributed to the risk of developing SSc, an Australian study found the relative risk (RR) to be increased in first degree relatives (RR 14.3) and siblings (RR 19.3) (31). An American study that included 703 families, eleven of which had multiple members affected reported similar results. The relative risk was found to increase from 13 in first-degree relatives to 15 for siblings (32). To date, this remains the highest risk factor for SSc development ever reported. However, it has to be noted that the absolute risk remains rather low (~1%). Based on these findings one may hypothesize that environmental factors play a role as well, but there seemed to be no socioeconomic or geographic clustering present in these populations. The genetic background of SSc patients is to some extent similar to the background underlying other autoimmune diseases (AID), which was recently illustrated by a study investigating the occurrence of autoimmune disease in first degree relatives of SSc patients. Of 719 SSc patients, 260 (36%) patients had first-degree relatives with at least one autoimmune disease, most notably RA (18%) and autoimmune thyroid disease (9%). No significant differences in the type and frequency of familial autoimmunity were observed between the diffuse and limited cutaneous subsets of disease (33). It has to be noted that, next to the age of the twins used in this study, other factors might play a role in the discrepancy between risks in twins and first-degree relatives. The most obvious explanation is probably that some environmental factors might play a role, but that they require a certain genetic background to act on, which unfortunately makes these factors hard to identify without adjusting for the genetic context of these individuals. Another explanation is that epigenetics play a role on determining the final clinical phenotype of having SSc, another AID or being healthy. For example, investigating methylation patterns in SSc patients, and especially in discordant twins might be very useful to determine which parts of the genome are important for determining the final phenotype. Methylation is influenced by environmental factors, and therefore elucidating methylation patterns in SSc and comparing them to explorations of the epigenome in other diseases might be helpful in determining the environmental

factors contributing to SSc.

Several case reports describing twins concordant for SSc have been published over the past decades, but it was not before 2003 that a large twin study was published (34,35). This study included 42 twin pairs (24 monozygotic and 18 dizygotic) of which at least one twin was suffering from SSc. Of the total 42 twin pairs only two pairs, both female-female, were concordant for SSc. One pair was dizygotic (DZ) and suffered from the limited cutaneous subform of the disease (lcSSc), the other pair was monozygotic (MZ) and one twin had lcSSc whereas the other had dcSSc. These data did not show a significant difference in concordance rate between MZ twins (4.2%) and dizygotic twins (5.6%). The overall cross-sectional concordance rate was 4.7%. (36). This is in contrast to the family studies from which it was clearly indicated that genetic factors play a role in the development of the disease. Interestingly, there was a significantly higher concordance for anti nuclear antibody (ANA) positive titers (≥1:40) in the MZ twins (90%) compared to the DZ twins (40%). MZ twins had a higher concordance rate for the presence of either ANAs or anticytoplasmic autoantibodies (ACA) compared to the DZ twins, respectively 95% and 60%. The mean age of the participants in this study was 48.5 years, ranging from 28 to 69 years. The most common age of onset for SSc is somewhere around 50 years of age, and for that reason a longer follow-up period might have been more appropriate to investigate the concordance rates in these twin pairs, the real concordance might thus be underestimated (7). ANA and ACA antibody titers are often present before onset of clinical disease in SSc, but clinical disease does not always develop when they are present (37). These observations might explain the markedly increased concordance for antibody titers but not for SSc. Zhou et al collected total RNA from dermal fibroblasts of 15 discordant twin pairs (10 MZ and 5 DZ) and healthy controls. This RNA was used in a microarray expression analysis. SSc fibroblasts had RNA expression profiles that were significantly different from those of unaffected DZ twins and normal controls. Of interest, unaffected MZ twins were not significantly different from SSc patients. Moreover, normal fibroblasts incubated with serum from a SSc patient or with serum from her unaffected MZ twin sister developed an expression pattern very similar to that observed in SSc fibroblasts. Based on these

observations the authors draw the conclusion that a genetic predisposition to SSc seems to be apparent at the molecular level in skin fibroblasts (38). This finding illustrates that, in respect to studying fibroblasts in SSc, it has to be taken into account that the genetic background of the fibroblast seems to be less important in determining disease compared to the molecules present in the serum of genetically similar individuals. This positions the fibroblasts downstream in the chain of events causing disease. On the other hand, it remains enigmatic why the fibroblasts of the healthy part of the MZ twin have similar RNA expression levels, without having SSc. This merely indicates that there needs to be an additional, possible environmental, factor leading to full blown disease and rather justifies the existence of the genetically “susceptible fibroblast”. These interesting findings need further investigations and a second large twin study would be of high interest in the field of SSc research.

Chromosomal studies in systemic sclerosis

Chromosomal instability seems to be present in a proportion of SSc patients, for instance findings of fragmented, dicentric and ring chromosomes have been reported, and more rarely, translocation and deletion of chromosome fragments, microchromosomes and double minute chromosomes (39-42). These findings are not only present in full blown SSc, but already in patients that have Raynaud’s phenomenon and will proceed to SSc later on (42). The current search for chromosomal aberrances in SSc is segregated in three directions. The first field of interest is telomere biology, provoked by observations of chromosomal aberrances in SSc. The second field focuses on the role of the X chromosome, prompted by the observation that SSc is far more frequently present in females. The third field focuses on physical aberrances that might trigger ANA (e.g. in centromeres). The latter will not be discussed since it leaves the realm of genetics and enters the field of functional research.

Telomeres are nucleoprotein complexes at the 5’and 3’end of the chromosome that play a pivotal role in DNA and, hence, chromosomal damage signalling and repair. The fundamental component of the telomere is a nucleotide repeat stretch that

shortens with cell division. An initial study investigating telomeres in systemic sclerosis took off from the observation of chromosomal instability in SSc and hypothesized that this instability resulted from a shortened telomeric repeat. The study investigated telomere lengths of peripheral blood leukocytes (PBLs) from 43 SSc patients, 182 SSc family and 96 age-matched controls. They found the average loss of telomeric DNA in SSc patients and their family members to be 3 kb compared to the controls. This loss occurred independent of age and disease duration. Of interest, this difference was not observed in fibroblasts from SSc patients, family and controls. The authors concluded that these results might reflect a genetic predisposition for shorter telomeres in these families, or exposure to a common environmental agent (43). A second study scrutinizing telomere shortening in SSc focused on the lcSSc phenotype only. This study investigated 43 lcSSc patients with age ranging from 37 to 80 years. The authors found telomere lengths in PBLs from lcSSc patients to be significantly longer than controls. In addition, shortening was not age-related and differed significantly from age-matched controls only after 50 years of age (44). Until now, literature addressing the role of telomeres in SSc is conflicting, but this may come forth from the different clinical subsets of SSc investigated by these studies. Noteworthy, a study investigating telomerase activity, which facilitates the maintenance of the telomere by adding repeat sequences to chromosomal ends, found a low telomerase activity in SSc compared to other autoimmune diseases and healthy controls (45). Taken these observations together, it is likely that aberrances in telomere biology are present in SSc, Chapter 11 of this thesis provides a further in depth study on the role of telomere shortening in SSc.

To level differences between males and females in X chromosomal gene expression, several species including mammals, evolved dosage compensation mechanisms. For instance, females randomly inactivate one of the two X chromosomes by methylation (46). As a consequence of this, females even-handedly inactivate the paternal and maternal X-chromosome throughout their cells (47). Although inactivation of the X chromosome is permanent for all descendants of a cell, the X chromosome inactivation (XCI) pattern alters with age. The frequency of skewed XCI, in other words a misbalance between the proportion of cells that inactivate the paternal X and the proportion that inactivate the maternal

X, increases in elderly compared to younger healthy females (48). Intriguingly, women with SSc comprise a significantly higher frequency of peripheral blood cells with a skewed XCI pattern compared to healthy women. Two overlapping studies from the same Turkish research group showed that in 195 female SSc patients and 160 female controls skewed XCI patterns were significantly more present; 44.9% of 149 informative patients and in 8% of 124 healthy controls. In addition, they found extremely skewed patterns to be more frequent in the SSc population (49, 50). The study presented in Chapter 10 of this thesis validates these results and describes the role of XCI in multiple immune cell subsets of SSc patients and its effects on *forkhead box P3* (*Foxp3*) expression, which is located on the X chromosome. (51).

X chromosome monosomy is another model proposed to play a role in the increased prevalence of SSc in females. It is a well-known fact that females with Turner's syndrome, harbouring only one X chromosome, have a vastly increased risk for developing autoimmune disease. A study investigating the presence of X monosomy in peripheral blood leukocytes from 44 females with SSc and 73 age-matched healthy women found that the monosomy rates in SSc were significantly higher compared with healthy women. No significant difference in monosomy rates was observed in patients with different types of SSc. Monosomy rates seemed to increase with age and were higher in B and T cells compared to monocytes/ macrophages, polymorphonuclear, or natural killer cells. The authors excluded the possibility of male cell microchimerism by ruling out the presence of the Y chromosome (52). Taken these observations together, there is cumulative evidence for a role of the X chromosome in SSc, which might bring more insights in the clustering of this disease in females.

The Human Leukocyte Antigen (HLA) region in systemic sclerosis

Many autoimmune diseases have been associated with variants in the HLA region and have inspired researchers to investigate this region in SSc. The HLA region remains an extremely polymorphic region in the genome, vastly influenced by ethnic origin, and might therefore play a role in the differences of SSc prevalence in diverse ethnic populations. The first endeavors to implement the HLA region in SSc susceptibility resulted in weak associations of SSc with a number of *HLA* alleles or

haplotypes. These concerned *HLA-DR5 haplotypes (DRB1*1101* and *1104*, *DQA1*0501*, *DQB1*0301*) and *DR3 (DRB1*0301, DQA1*0501, DQB1*0201)* in American and European white populations and *HLA-DR2 (DRB1*1502, DQB1*0601)* in Japanese (53-55). In the Choctaw Native Americans, the variants in *HLA-DR2 (DRB1*1602, DQA1*0501, DQB1*0301)* were found to pose an increased risk (56). A study of 98 patients with SSc from Greece, found *HLA-DRB1*1104*, in linkage disequilibrium with *DQA1*0501, DQB1*0301*, associated with SSc susceptibility and ATA but not ACA antibodies (57). The significance levels of HLA associations, however, increase when SSc patients are subcategorized according to their specific autoantibody subsets. These included associations with the characteristic SSc auto-antibodies. (55, 56, 58). In 2010 two large studies were published with complementary findings. The first investigated 1300 SSc cases (961 white, 178 black and 161 Hispanic) with extensive autoantibody characterization (ACA, ATA, anti-RNA polymerase III (ARA), and fibrillarin) to investigate predisposing effects of the *HLA* alleles. They found the strongest associations with SSc in white and Hispanic SSc patients (*DRB1*1104*, *DQA1*0501, DQB1*0301* haplotype and *DQB1* alleles) and a protective effect for both lcSSC and dcSSc of *DRB1*0701, DQA1*0201*, *DQB1*0202* haplotype and *DRB1*1501* haplotype, subsequently in a dominant and recessive model. SSc in black subjects was associated with *DRB1*0804*, *DQA1*0501* and *DQB1*0301* alleles. Anti-topoisomerase antibodies were associated with *DPB1*1301*. Associations with ACA (*DQB1*0501* and *DQB1*26*) and ARA (Caucasian and Hispanic: *DRB1*0404*, *DRB1*11* and *DQB1*03*, Black patients: *DRB1*08*) were also described (59). A Spanish study in 100 patients postulated very similar findings and found the *HLA-DRB1*11* to be associated with SSc susceptibility and the *HLA-DRB1*0701* being protective as well (60). In the American cohort described above, a novel strategy to identify variation in the HLA region was applied (61). This sequence feature variant type approach replicated the previous association between SSc and *HLA-DRB1*1104* in Caucasians and Hispanics. The most interesting part of this study is that they found a common amino-acid make up of *DRB1*1104* and *HLA-DRB1*0804*, associated in Caucasians and Hispanics. This finding might help explain why two different DRB1 alleles are both associated with SSc but in different races.

More recently, two small-scale studies appeared that implicated HLA alleles in

development of complications during disease course. The first described a 3-factor model comprising two single nucleotide polymorphisms (SNPs), *IL-6 174C>G* and the *IL2 330G>T*, and the *HLA-B*3501* allele that was predictive for the occurrence of digital ulcers in 200 Italian SSc patients (62). The second study describes 90 patients, derived from a cohort of 1,519 SSc patients that developed scleroderma renal crisis. This study found that *DRB1*0407* and **1304* are risk factors for the development of scleroderma renal crisis, independent of previously postulated risk factors (63).

Candidate gene studies

Several studies investigate genes that seem to play a role in SSc susceptibility. Selection of candidates is often based on either the functional implication of the gene in pathogenesis, or coming forth from studies in similar diseases. In SSc both have been done, although most targets for candidate gene or variant studies are derived from reports of associations in other AID. Association studies most often make use of the presence of variants or polymorphisms in genes. In the studies described below, single nucleotide polymorphisms were investigated. These are best described as single basepair transitions (e.g. a change from nucleotide A to G at a certain genomic locus). Based on their position in or around a gene they can have different impact on gene function. When the SNP is present in an exon, it might cause an amino-acid change or transcription stop at that position. Presence in a splice region might confer to alternative splicing and presence in the promoter or 3'UTR region to a change of gene expression. SNPs can also be used as indicator for the genomic region nearby their own locus; this principle is based on linkage disequilibrium. Therefore, when a SNP is found associated with a certain disease, the conclusion cannot unequivocally be that this SNP itself is contributing to disease susceptibility, but it is likely that a real causative variant can be found in linkage with this SNP. Both features of SNPs described above are used for association studies in SSc. Some studies have selected variants ultimately causing a change in protein function for their candidate gene studies. Other studies have selected those SNPs that provide information of a large part of the gene of interest based on linkage disequilibrium. Chapters 2 to 8 describe polymorphisms that are investigated for a possible role in SSc susceptibility and pathogenesis. Next to this,

polymorphisms investigated in Chapters 6 to 8 are assessed for a possible functional impact on gene expression or inflammatory potential. Polymorphisms described in this section are best regarded as risk factors for developing SSc or related clinical complications, not as diagnostic tools. Below an overview of studies with sufficient patients and power to be reproducible and, hence, able to ascertain or exclude involvement of the selected candidate variants in SSc. The last paragraphs of this section will discuss current unclear and conflicting findings. **Table 4** provides an overview of all large genetic association studies performed in SSc so far.

Table 4. Established non-HLA susceptibility genes for SSc and its clinical phenotypes

Gene	SNP (associated allele)	With	Population	Size (tested phenotype/ HC)	OR (95%CI)	Reference
STAT4	rs7574865 (T allele)	lcSSc	European	1317/3113	1.54 (1.36-1.74)	65
		lcSSc	Japanese	282/590	1.35 (1.10-1.66)	66
		ACA+	Japanese	87/590	1.62 (1.17-2.22)	66
		SSc	French	885/970	1.29 (1.11-1.51)	67
	rs11889341 (dominant A allele)	SSc	North-American	902/4745	1.29 (1.20-1.50)	68
IRF5	rs2004640 (TT genotype)	SSc	French	881/760	1.58 (1.18-2.11)	72
		ANA+	French	608/760	1.59 (1.16-2.17)	72
		Fibrosing alveolitis	French	280/760	2.07 (1.38-3.11)	72
		SSc	Japanese	281/477	1.23 (0.91-1.65)	73
		dcSSc	Japanese	142/477	1.27 (0.87-1.85)	73
		ATA+	Japanese	87/477	1.56 (0.98-2.49)	73
	rs10954213 (AA genotype)	SSc	Japanese	281/477	1.41 (1.0.1-1.96)	73
		dcSSc	Japanese	142/477	1.38 (0.90-2.10)	73
	rs2280714 (AA genotype)	SSc	Japanese	281/477	1.72 (1.18-2.50)	73
		dcSSc	Japanese	142/477	2.05 (1.23-3.41)	73
		ATA+	Japanese	87/477	2.10 (1.12-3.95)	73
TBX21	rs11650354 (TT genotype)	SSc	North-American	902/4745	3.37 (2.40-4.60)	79

Gene	SNP (associated allele)	With	Population	Size (tested phenotype/ HC)	OR (95%CI)	Reference
Bank1	rs10516487 (G allele)	SSc	European and North American	2362/3248	1.12 (1.03-1.22)	83
		dcSSc	European and North American	770/3284	1.20 (1.05-1.37)	83
		ATA+	European and North American	435/3284	1.20 (1.02-1.41)	83
	rs10516487 (T allele)	dcSSc	European	431/1118	0.77 (0.64-0.93)	83
	rs17266594 (T allele)	SSc	European and North American	2351/3231	1.14 (1.05-1.25)	83
		dcSSc	European and North American	613/3231	1.23 (1.08-1.41)	83
		ATA+	European and North American	478/3231	1.24 (1.05-1.46)	83
	rs3733197 (G allele)	dcSSc	European and North American	613/3143	1.15 (1.02-1.31)	83
		ATA+	European and North American	447/3143	1.26 (1.07-1.47)	83
	rs3733197 (A allele)	dcSSc	European	431/1137	0.73 (0.61-0.87)	84
BLK	rs2736340 (TT genotype)	SSc	European and North American	1639/1416	1.71 (1.2-2.4)	86
		ACA+	European and North American	510/1416	2.27 (1.5-3.5)	86
	rs2736340 (CT genotype)	SSc	European and North American	1639/1416	1.31 (1.1-1.5)	86
		ACA+	European and North American	510/1416	1.6 (1.2-2.0)	86
	rs13277113 (AA genotype)	SSc	European and North American	1639/1416	1.25 (1.1.-1.4)	86
		ACA+	European and North American	510/1416	1.42 (1.2-1.7)	86
	rs13277113 (A allele)	SSc	Japanese	309/769	1.45 (1.17-1.79)	87
TNFSF4	rs2205960 (T allele)	SSc	North-American	1059/698	1.2 (1.1-1.5)	89
		ATA+	North-American	174/698	1.4 (1.1-1.9)	89
	rs1234314 (G allele)	SSc	North-American	1059/698	1.2 (1.04-1.4)	89
		ACA+	North-American	300/698	1.3 (1.1-1.6)	89
		ATA+	North-American	174/698	1.3 (1.02-1.7)	89
		SSc	European	2856/2920	1.15 (1.02-1.31)	90
		lcSSc	European	1608/2920	1.22 (1.07-1.38)	90
		ACA+	European	828/2920	1.23 (1.10-1.37)	90
	rs844648 (A allele)	SSc	North-American	1059/698	0.8 (0.7-0.97)	89
		ARA+	North-American	193/698	1.4 (1.1-1.8)	89
		lcSSc	European	1673/2977	1.1 (1.01-1.20)	90
		ACA+	European	860/2977	1.12 (1.01-1.25)	90
	rs844644 (A allele)	lcSSc	European	1653/2946	0.91 (0.83-0.99)	90
		ACA+	European	856/2912	0.90 (0.80-1.00)	90

Gene	SNP (associated allele)	With	Population	Size (tested phenotype/ HC)	OR (95%CI)	Reference
	rs12039904 (T allele)	SSc	European	2894/2991	1.18 (1.08-1.29)	90
		lcSSc	European	1639/2991	1.20 (1.09-1.33)	90
		ACA+	European	840/2991	1.22 (1.07-1.38)	90

SSc: Systemic Sclerosis, lcSSc: limited cutaneous systemic sclerosis: dcSSc: diffuse cutaneous systemic sclerosis, ACA: anti-centromere antibodies, ATA: anti-topoisomerase antibodies, ARA: anti-RNA polymerase antibodies.

Signal transducer and activator of transcriptor 4 (STAT4) has been associated with rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE) (64). This prompted a study, described in Chapter 2 of this thesis, in five independent European cohorts totaling 1317 SSc patients and 3113 healthy controls. This initial study found the *STAT4* rs7574865 T allele associated with lcSSc but not with dcSSc (65). The very same association was later on confirmed in a Japanese cohort, which also added an association with ACA positivity (66). Subsequently, two reports replicated the association of *STAT4* with SSc susceptibility but in contrast, observed the association in both lcSSc and dcSSc (67, 68). These studies together strongly implicated *STAT4* in SSc susceptibility. Interestingly, STAT4 plays a role in interleukin (IL)12 and IL23 receptor signalling and seems to be activated through interferon (IFN) type 1 receptors. STAT4 functions as a T helper cell (Th) 1 promoter (69). A recently developed murine model investigated the susceptibility for bleomycine-induced fibrosis in *STAT4* knockout mice. This model showed that mice deficient for STAT4 had decreased T cell activation, proliferation and cytokine release, which led to less fibrosis. Creating *STAT4* knockout Tsk1 mice (which develop extensive fibrosis and tightening of the skin during life) however, did not ameliorate their fibrotic phenotype (70).

Interferon regulatory factor 5 (IRF5) polymorphisms have been associated with SLE (71). In SSc this gene also seems to play a role. A French study investigated 427 patients with SSc and 380 healthy controls and a replication set comprising 454 patients with SSc and 380 control subjects (72). This study found that the TT genotype of the functional polymorphism rs2004640 in *IRF5* was significantly enriched in patients with SSc compared to healthy controls. Analyses in all SSc patients revealed a significant association between homozygosity for the T allele and the presence of ANA and fibrosing alveolitis. This study was confirmed by a

study in Japanese SSc patients (73). Interestingly, a follow up on the French association study was published and showed that haplotypes of the *IRF5* gene were more informative in predicting dcSSc and fibrosing alveolitis (74). IRF5 is able to induce IFN alpha related genes (75), increased IFN alpha levels produced by plasmacytoid dendritic cells have previously been described in SSc, implicating the importance of this pathway on a functional level as well (76).

T-box expressed in T cells (T-bet, TBX21) knockout mice develop extensive fibrosis upon injection with bleomycin (77, 78). In addition, TBX21 is an important transcription factor involved in Th1 development and seems to play a role in B cell and dendritic cell biology as well. Since this transcription factor appears to be intricately involved in the process of fibrosis, this gene was of interest to investigate in SSc. A study in American Caucasian patients and controls pointed out that the variant rs11650354 of *TBX21* was involved in SSc using a recessive model. The association was present regardless of clinical subtype or autoantibody pattern. Furthermore, this polymorphism seemed to be functional because SSc patients homozygous for the rs11650354 TT variant expressed elevated Th2 cytokines, whereas carriers of the CC genotype were characterized by a type I interferon signature (79).

B-cell scaffold protein with ankryn repeats (BANK1) is expressed in B cells and is tyrosine phosphorylated upon B-cell antigen receptor stimulation. Overexpression of BANK in B cells leads to enhancement of BCR-induced calcium mobilization (80). *BANK1* polymorphisms have been previously associated with SLE and RA and have now been implicated in SSc susceptibility as well (80-83). The initial large multi-centre study, encompassing 2380 white SSc patients and 3270 matched controls, found a strong association with the rs10516487 G and rs17266594 T alleles and the diffuse cutaneous subform of the disease. This observation was further strengthened by a combined French and German effort, which also found these alleles to be strongly associated with the dcSSc subform (84). However, the functional implication of these polymorphisms still remains unclear.

B lymphoid kinase (Blk) is another susceptibility locus coming forth from genetic studies in SLE. More specifically it concerns the *FAM167A-BLK* region, previously

entitled C8orf13-BLK. BLK is a kinase that is expressed in thymocytes and involved in downstream signalling of the B cell receptor (85). Until now, two studies found an association of this region with both lcSSc and dcSSc. Both studies investigated two polymorphisms, rs13277113 and rs2736340. The T variant of rs2736340 was associated with SSc in both an U.S. and a Spanish case-control population. The A variant of rs13277113 was only associated with SSc in the U.S. cohort. Both variants demonstrated an association with ACA and lcSSc in a combined analysis. RNA expression profiling using peripheral blood pointed out different profiles of BCR related pathways when patients were clustered to their genotype (86). A Japanese study confirmed the association of rs13277113 with SSc, but did not investigate rs2736340 (87).

Tumor necrosis factor superfamily-4 (TNFSF4 or OX40 ligand (OX40L)) is the ligand for OX40 present on T cells. OX40 ligation has been shown to enhance T-cell expansion and survival and in addition, decreases the suppressive activity of regulatory T cells (88). *TNFSF4* polymorphisms have been associated with SLE susceptibility. A cohort consisting of 1059 Caucasian SSc cases and 698 healthy controls revealed a significant association between susceptibility to SSc and the minor alleles of SNPs rs1234314, rs2205960 and rs844648. Additional analysis of clinical phenotypes demonstrated significant associations of *OX40L* variants with lcSSc and dcSSc as well as anti-topoisomerase (ATA), ACA and POL-positive patients (89). A replication study, displayed in Chapter 5, confirmed the influence of *TNFSF4* polymorphisms in SSc susceptibility, most marked in subsets of patients with lcSSc and ACA and this study also identified new risk haplotypes involved in SSc susceptibility (90).

Next to these confirmed, associated genes there are a number of genes that are very interesting but found ambivalently involved in SSc when searching the published literature. These ambivalent findings might be coming forth from environmental differences, differences in ethnic background or might be caused by a type I or II error. These candidate genes are described below.

Interleukin 23 (IL23) promotes proliferation of Th17 cells. It has been postulated that Th17 cells play a central role in SSc pathogenesis (91). In contrast to multiple

other autoimmune diseases, two studies showed that *IL23R* variants do not confer risk to SSc susceptibility (92, 93). This was recently nuanced by an American study, which did not observe an association with SSc susceptibility as well, but nevertheless found that the IL-23R variants rs11209026 and rs11465804 were associated with ATA and pulmonary arterial hypertension (94). Although apparent discrepancies between the studies might be coming forth from differences in subgroup analysis strategy, further investigations are necessary before *IL23R* can be added to the list of SSc susceptibility genes.

The *FAS* gene has been described as an “autogene”, because its dysregulated functioning contributes to various autoimmune diseases. One of the main activators of apoptosis in T cells is soluble FAS; which has been found elevated in SSc serum (95, 96). A recessive model of the *FAS* -670 GG genotype revealed a strong association with SSc, lcSSc and ACA+ lcSSc in a multinational cohort consisting of 2900 SSc and 3186 healthy controls (97). These findings, displayed in Chapter 3, are in contradiction with conclusions previously reported. In an Italian study, encompassing 350 SSc patients and 232 healthy controls, the other allele of *FAS -670* was found to be associated with SSc susceptibility. As an underlying factor for this apparent contradiction it is noteworthy to mention that the minor allele frequency in the Italian control group differs from the allele frequency reported by the international HAPMAP project (www.hapmap.org) and other Italian studies investigating the *FAS* -670 A>G polymorphism (98).

Connective tissue growth factor (CTGF) has been implicated in SSc on a functional level. Several studies show CTGF up regulated in SSc skin and CTGF is able to promote fibrosis by enhancing extracellular matrix deposition and proliferation of mesenchymal cells. These observations put forward the hypothesis that *CTGF* polymorphisms might play a role in SSc (99). The first study describing an association of the homozygous GG genotype of a variant in *CTGF* (-945G>C) was performed in 500 SSc patients and 500 controls. This report also shows that the substitution of cytosine for guanine creates a binding site for the transcription factors Sp1 and Sp3. The *CTGF*-945 C allele has high affinity for Sp3 and reduces transcriptional activity. This was functionally confirmed with an immunoprecipitation assay (100). A Japanese report substantiated these findings, but two large studies

in European and North-American cohort were not able to replicate the findings (101-103). Of relevance is that nearly 50% of the Japanese cohort were dcSSc patients, with typical clinical hallmarks as ATA and interstitial lung disease. These were exactly the SSc patients subsets that had a higher frequency of the homozygous GG genotype in *CTGF*. On the other hand, a later author reply of the first UK study indicates that they used the nomenclature for the *CTGF* polymorphism as it was first described, stating that the *CTGF*-945GG genotype in their study is the same as the *CTGF*-945CC variant in the other studies and vice versa, because nomenclature was based on the complementary DNA strand. The Japanese study does not indicate which nomenclature was used, depending on this, the results might indeed be substantiated or in fact be the total opposite. Interestingly, a more recent French study describes a transcription altering *CTGF* polymorphism (rs9399005), located in the 3'UTR region of *CTGF* having an protective effect for SSc development. Taken these observations together, the exact role of *CTGF* polymorphisms in SSc remains a matter of debate and still warrants more research (104).

In the past few months three additional reports have been published, giving new insights in the genetic background of SSc. These are *PTPN22*, *UPAR (CD87)*, *CD226* and *NLRP1*. Although these genes seem to be related to the pathogenesis of SSc and its subtypes, future replication studies are necessary to establish these genes as risk factors implicated in SSc susceptibility (105-108).

Genome-wide association studies

To date, only a few studies have investigated the genetic make-up of SSc utilizing a genome-wide approach. Although their non-hypothesis based approach is the same, marked differences in robustness of the methodology used resulted in deviating results.

The first study appeared in 2003 and described a genome-wide microsatellite screen at 10 cM resolution, including 400 markers in 20 Choctaw patients with SSc and 76 matched controls. Based on the results of the initial screen, fine- mapping at < or =1 cM resolution was performed. From the genome-wide screen, 17

markers were associated with SSc in this population. The novel regions linked to SSc susceptibility were *1p32-31, 7q35, 8q24.12, 19p13.2, 22q13.1* and *Xq21-23*. These included loci close to the *SPARC*, *MHC*, *FBN1*, and *TOPOI* genes (109).

A second genome-wide association study was performed in 137 Korean SSc patients and 564 controls using 500,568 SNP markers. This genome scan showed a significant peak located in the region of *HLA–DPB1* and *DPB2* on chromosome 6p. The peak included the SNPs rs3128930, rs7763822, rs7764491, rs3117230, and rs3128965. Fine-mapping of this region revealed that rs3128930, rs7763822, rs7764491, rs3117230, and rs3128965 were the SNPs associated with SSc. The association was stronger in patients with ATA. These SNPs also showed an association with dcSSc, but not with lcSSc. After fine mapping, the results were validated in 1,107 Caucasian US SSc patients and 2,747 controls. In this replication cohort two pairs of SNPs, rs7763822/rs7764491 and rs3117230/ rs3128965, showed an association with SSc in patients who had either circulating ATA or ACA, but the association with SSc susceptibility overall was not replicated. Interestingly, SNPs corresponding to genes that have previously been associated with SSc, such as *PTPN22* and *CTGF* showed low p-values (10^{-5}–10^{-6}) but failed to reach genome-wide significance (110).

In 2010 however, a large and robust genome-wide association study was published. This study included 2296 SSc patients and 5171 healthy controls originating from The Netherlands, Germany, Spain and the United States. Different sets of markers were used that ranged in density from 308,349 to 488,793 SNPs. This study adjusted for multiple testing by regarding the genome-wide significance threshold ($P \leq 5\times10^{-7}$) as their significance threshold. The strongest association was observed at the 6p21 locus in the midst of the MHC region. At this region, rs6457617, situated in *HLA*DQB1*, showed the strongest association. In addition, five non-HLA loci showed genome-wide significance; *TNPO3/IRF5* region in *7q32, STAT4* in *2q32*, *CD247* in *1q22-23*, *CDH7* in *18q22* and *EXOC2/IRF4* near *6p25*. These results further establish *STAT4* and *IRF5* as genetic risk factors for SSc. To verify the novel findings, a case-control set comprising 2753 SSc patients and 4569 controls were genotyped for the three, not previously described variants. Two SNPs identified in the discovery cohort in the *EXOC2/IRF4* and *CDH7* regions

were not replicated in the validation cohort. However, the initial association with rs2056626 in the *CD247* gene was replicated (111). *CD247* is of particular interest, considering it has an important role in the immune system by encoding the T-cell receptor zeta subunit, which is a component of the T cell receptor complex, this gene has also been associated with susceptibility to SLE previously (112). The association of *CD247* was recently independently replicated in a French Caucasian population consisting of 1031 SSc patients and 1014 healthy controls (113).

A follow up study, exploiting the above mentioned GWAS data, focused on determining genetic components contributing to lcSSc, dcSSc, ACA+ and ATA+. To this purpose, a meta-analysis was performed in four cohorts, comprising 2,296 SSc patients and 5,171 healthy controls. Subsequently, 18 polymorphisms with a P value lower than $1x10^{-5}$, seven in the lcSSc subtype, five in the dcSSc subtype, two in ACA positives and four in ATA positives were further tested in nine independent cohorts composed of an additional 3,175 SSc patients and 4,971 controls. Overall analysis revealed one variant in the *interferon regulatory factor 8 (IRF8)* gene (rs11642873) to be associated with lcSSc at genome wide significance. Variants in the *GRB10 growth factor receptor-bound protein 10 (GRB10)* and sex determining region *Y-box 5 (SOX5)* were just below the genome wide significance threshold associated with, respectively, lcSSc and ACA+ subgroups. Intriguingly the authors propose a model of *IRF8* and *SOX5* affecting the formation of extra-cellular matrix through collagen, type II, alpha 1 in the skin and other organs of SSc patients. Furthermore this study revealed genome wide significant results in the *HLA-DQB1* locus for ACA+ patients (OR = 2.48), in the *HLA-DPA1/B1* loci for ATA+ patients (OR = 8.84) and in *NOTCH4* for ACA+ patients (OR = 0.55) and ATA+ (OR = 0.54). This study is of particular interest since it indicates that the heterogeneity of SSc phenotypes is likely to be the reflection of a different genetic foundation. (114)

Very recently, the fourth GWAS was published which used a high-resolution marker set comparable to the previous study (approximately 500,000 markers). The study used a two-step approach; the first step consisted of a GWAS and was conducted in a French cohort consisting of 564 cases and 1,776 controls. Although only one SNP showed genome wide significance, which was again in the *HLA*DQB1* region, the authors validated 20 SNPs (the 17 top significant SNPs and 3 previously

associated SNPs) in a replication cohort of 1,682 SSc cases and 3,926 controls. Follow-up of the 17 top SNPs revealed associations at *PSORS1C1* (HLA region), *TNFAIP3 interacting protein 1 (TNIP1)* and *ras homolog gene family, member B (RHOB)* loci. Furthermore, the associations of previously identified candidate loci *STAT4*, *IRF5* and *CD247* were substantiated. This study furthermore addressed the functional relevance of their findings by investigating *TNIP1* gene and protein expression. TNIP1 was expressed at a lower level in both SSc lesional skin tissue and cultured dermal fibroblasts from SSc patients based on genotype. Intriguingly, *TNIP1* showed in vitro inhibitory effects on cytokine-induced collagen production. Although this study presents novel and interesting candidate loci, it has to be noted that only one locus reached genome wide significance initially. The other loci become highly significant after pooling of the first and second step, but are strictly speaking not significant yet in the initial French GWAS. These loci not being identified from the first step may be a power problem coming forth from the relatively low number of patients included in the GWA step. In addition, the study does not describe if there is an overlap between cohorts used for previous studies identifying SSc susceptibility genes *STAT4*, *IRF5* and *CD247* and the ones used in this study for replication, which would be welcome to place these results in the correct perspective (115).

Gene-gene interaction studies

As discussed in the previous sections, a large number of genetic variants seems to be involved in SSc susceptibility and pathogenesis. These associated genes all have a very modest, but reproducible effect on susceptibility throughout populations. In line with the paradigm of multifactorial diseases, one would expect to be at higher risk for SSc when harboring more susceptibility variants. Not until recently, genetic research in the field of SSc started to combine genetic data to see whether some variants together formed an additive risk for SSc susceptibility. An additional thought that justifies these attempts is the fact that many SSc candidate genes map to the same biological pathways (116).

The first successful attempt showed that *STAT4* (rs7574865) and *IRF5* (rs2004640) variants form an additive risk for development of SSc and interstitial lung disease

(67). After this first study, this research group repeated the analysis including the *BANK1* polymorphisms and could display an additive effect with regard to diffuse SSc susceptibility (84). In a subsequent analysis they added a *NLRP1* polymorphism to the list of variants interacting with *STAT4* and *IRF5* (117). In contrast, the SSc-associated *BLK* region had an additive effect with *BANK1* in the dcSSc subset (84). An American study showed that the *STAT4* polymorphism was predominantly enriched in SSc patients who carried the 'CC' genotype at *TBX21* rs11650354 (68). Next to this, an interesting study was published last year, describing a 3-factor model comprising two single nucleotide polymorphisms (SNPs): *IL6* -174C>G and the *IL2* -330G>T, and the *HLA-B*3501* allele that was predictive for the occurrence of digital ulcers in 200 Italian SSc patients (62).

These initial studies on interactions of susceptibility genes look promising, but they still display only a small proportion of the risk for development of SSc. Chapter 9 provides a novel method to investigate epistatic interactions between genes in SSc. The most promising findings of gene-gene interactions playing a role will probably come forward from genome wide interaction analysis. However, this kind of analysis is currently hampered by a lack of both computational power and robust statistical methods. It is intriguing, however, that most polymorphisms associated with SSc map to pathways involved in T cell function and cell-cell interaction, as displayed in **figure 3**. Underscoring the role of the immune system in SSc and the need to look at immune cell subsets when investigating the functional impact of polymorphisms (116).

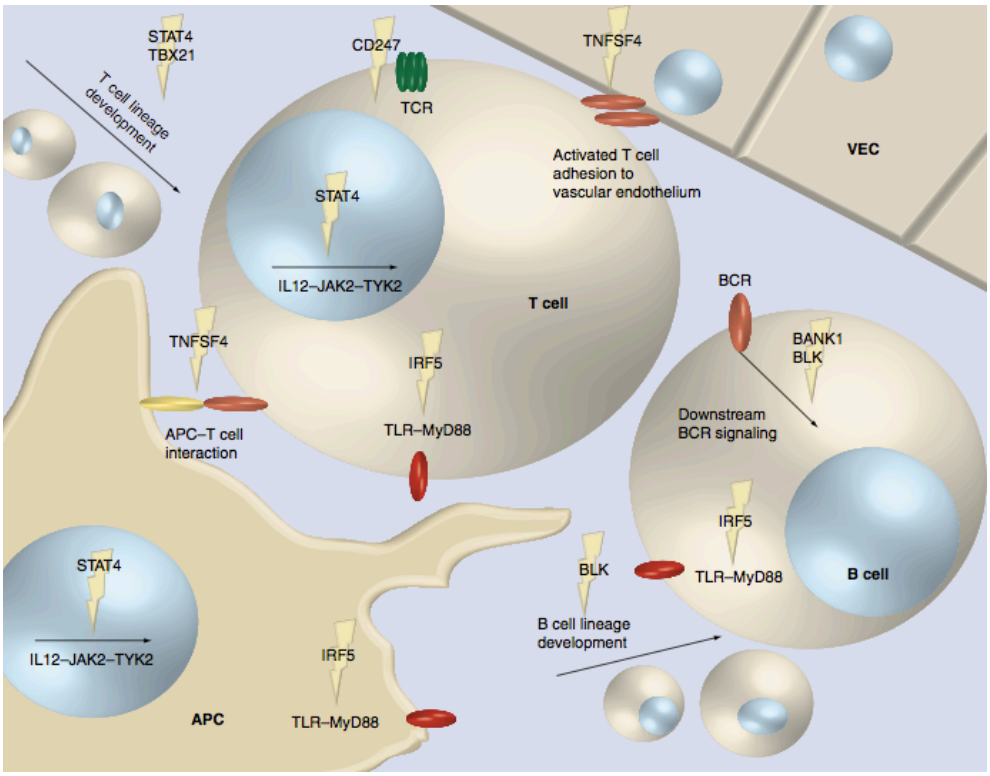


Figure 3. Graphical representation of current confirmed SSc candidate genes and their involvement in immunological processes (116). APC: antigen presenting cell; VEC; vascular endothelial cell, TLR: Toll like receptor, TYK: Tyrosine kinase, BCR: B cell receptor, TCR: T cell receptor JAK: Janus Kinase

Aim and outline of this thesis

The aim of this thesis is to decipher a part of the genetic background of systemic sclerosis and its clinical phenotypes. Genetic research in multifactorial disease has been criticized due to the uncertainty of its clinical and functional relevance, therefore this thesis targets at investigating the impact of polymorphisms on functioning of the immune system. When possible, the investigated polymorphisms will be reviewed in the light of development of clinical complications in time. This makes them relevant not only for determining susceptibility, but also as predictors of symptom development. The aim of the last chapter is to bring epigenetic and ageing related phenomena closer to immunological changes observed in SSc.

Part I (Chapters 2-5)

The first part focusses on candidate genes that are likely to be involved in SSc susceptibility, based on their associations with other autoimmune diseases or possible impact on disease pathogenesis. It comprehends a study describing a strong association of a *STAT4* polymorphism on SSc susceptibility, a variant previously associated with multiple AID. In addition, evidence is provided that a functional polymorphism in the apoptosis regulating gene *FAS* is implicated in the susceptibility for ACA+ lcSSc. Chapter 5 consists of a large replication study, firmly establishing *TNFSF4* polymorphisms in SSc susceptibility. On the other hand, Chapter 4 shows that a polymorphism in the *CD89* gene, previously associated with SLE, is not a common risk factor for SSc and RA.

Part II (Chapters 6-9)

The second part aims to combine both functional and genetic research to unveil the role of polymorphisms in the (dys)function of immune cells in SSc and development of clinical complications. It has been suggested that Interleukins 4 and 13 play a role in SSc. In Chapter 6 we investigated functional polymorphisms of these genes for their role in SSc and related clinical complications. In addition, we investigated whether these polymorphisms affected the expression of IL4 and IL13 in various cells of the immune system. Overall, these polymorphisms do not seem to play a major role in SSc. Chapter 7 consists of a study on the role of functional *TLR* variants in SSc. The study involves three steps, in the first, various *TLR* polymorphisms are investigated in a discovery cohort. A polymorphism in the *TLR2* gene is found to be associated with the diffuse subform of SSc, which is replicated in a large validation cohort. On a cellular level this polymorphism augments the immune response of dendritic cells and clinically it seems to play a role in pulmonary complication development in the 15 years after disease onset. Chapter 8 illustrates the observation that PBEF/Visfatin levels are increased in SSc patients with PAH. We investigate whether two expression regulating polymorphisms are playing a role in the increased expression and subsequently in PAH development. The final chapter of Part 2 describes a method for investigating epistatic interactions, using a cohort of SSc patients as a proof of principle for a functional interaction.

Part III (Chapters 10 and 11)

The final part expands on genomic foundations towards epigenetic phenomena observed in SSc, enhanced telomere shortening and an increased frequency of X chromosomal inactivation in SSc females. Both are in healthy persons related to biological ageing. Intriguingly, telomere shortening in SSc seems to be immune cell subset specific and X chromosomal inactivation seems to lose its relation with ageing in SSc.

Taken together, the parts and chapters of this thesis are interconnected by their common goal to investigate the genetic background of SSc with hypotheses based on emerging clinical and immunological findings in SSc and AID. This thesis therefore closely follows changing opinions on the genetics research and pathogenesis of SSc. As a consequence of this, the focus in this thesis shifts from basic SNP association studies, to evaluating functional and clinical relevance of polymorphisms. After publication of a genome wide scan in SSc it became clear that changes at the DNA level do not fully explain the risk for developing SSc, Therefore the final focus of this thesis is aimed at epigenetics research. In total, these chapters converge in a common aim to decipher those genetic processes likely to influence SSc pathogenesis based on contemporary scientific views on SSc.

References

1. David M. A case of scleroderma mentioned by Hippocrates in his aphorisms. Korot. 1981;8(1-2):61-3.

2. Rodnan GP. Signorina Galieri's scleroderma or scleredema. Ann Intern Med. 1972 Apr;76(4):673-4.

3. GOETZ RH. The heart in generalized scleroderma; progressive systemic sclerosis. Angiology. 1951 Dec;2(6): 555-78.

4. Ranque B, Mouthon L. Geoepidemiology of systemic sclerosis.Autoimmun Rev. 2010 Mar;9(5):A311-8. Epub 2009 Nov 10.

5. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. Arthritis Rheum. 1980 May; 23(5):581-90.

6. Roberts-Thomson PJ, Jones M, Hakendorf P, Kencana Dharmapatni AA, Walker JG, MacFarlane JG, Smith MD, Ahern MJ. Scleroderma in South Australia: epidemiological observations of possible pathogenic significance. Intern Med J. 2001 May-Jun;31(4):220-9.

7. Mayes MD, Lacey JV Jr, Beebe-Dimmer J, Gillespie BW, Cooper B, Laing TJ, Schottenfeld D.Prevalence, incidence, survival, and disease characteristics of systemic sclerosis in a large US population.Arthritis Rheum. 2003 Aug; 48(8):2246-55.

8. Chifflot H, Fautrel B, Sordet C, Chatelus E, Sibilia J. Incidence and prevalence of systemic sclerosis: a systematic literature review. Semin Arthritis Rheum. 2008 Feb;37(4):223-35. Epub 2007 Aug 9.

9. Arias-Nuñez MC, Llorca J, Vazquez-Rodriguez TR, Gomez-Acebo I, Miranda-Filloy JA, Martin J, Gonzalez-Juanatey C, Gonzalez-Gay MA. Systemic sclerosis in northwestern Spain: a 19-year epidemiologic study. Medicine (Baltimore). 2008 Sep;87(5):272-80.

10. Tamaki T, Mori S, Takehara K.Epidemiological study of patients with systemic sclerosis in Tokyo. Arch Dermatol Res. 1991;283(6):366-71

11. Arnett FC, Howard RF, Tan F, Moulds JM, Bias WB, Durban E, Cameron HD, Paxton G, Hodge TJ, Weathers PE, Reveille JD. Increased prevalence of systemic sclerosis in a Native American tribe in Oklahoma. Association with an Amerindian HLA haplotype. Arthritis Rheum. 1996 Aug;39(8):1362-70

12. Oliver JE, Silman AJ.Why are women predisposed to autoimmune rheumatic diseases? Arthritis Res Ther. 2009;11(5): 252. Epub 2009 Oct 26.

13. Selmi C, Invernizzi P, Gershwin ME. The X chromosome and systemic sclerosis.Curr Opin Rheumatol. 2006 Nov;18(6):601-5.

14. Laing TJ, Gillespie BW, Toth MB, Mayes MD, Gallavan RH Jr, Burns CJ, Johanns JR, Cooper BC, Keroack BJ, Wasko MC, Lacey JV Jr, Schottenfeld D. Racial differences in scleroderma among women in Michigan. Arthritis Rheum. 1997 Apr;40(4):734-42.

15. Le Guern V, Mahr A, Mouthon L, Jeanneret D, Carzon M, Guillevin L. Prevalence of systemic sclerosis in a French multi-ethnic county. Rheumatology (Oxford). 2004 Sep;43(9): 1129-37. Epub 2004 Jun 22.

16. Klein-Weigel P, Opitz C, Riemekasten G. Systemic sclerosis - a systematic overview: part 1 - disease characteristics and classification, pathophysiologic concepts, and recommendations for diagnosis and surveillance. Vasa. 2011 Jan;40(1):6-19.

17. Opitz C, Klein-Weigel PF, Riemekasten G. Systemic sclerosis - a systematic overview: part 2 - immunosuppression, treatment of SSc-associated vasculopathy, and treatment of pulmonary arterial hypertension. Vasa. 2011 Jan;40(1):20-30.

18. Trojanowska M. Cellular and molecular aspects of vascular dysfunction in systemic sclerosis. Nat Rev Rheumatol. 2010 Aug;6(8):453-60. Epub 2010 Jun 29.

19. van Bon L, Cossu M, Radstake TR. An update on an immune system that goes

awry in systemic sclerosis. Curr Opin Rheumatol. 2011 Nov;23(6):505-10.

20. Lafyatis R, York M. Innate immunity and inflammation in systemic sclerosis. Curr Opin Rheumatol. 2009 Nov;21(6): 617-22.

21. Radstake TR, van Bon L, Broen J, Wenink M, Santegoets K, Deng Y, Hussaini A, Simms R, Cruikshank WW, Lafyatis R. Increased frequency and compromised function of T regulatory cells in systemic sclerosis (SSc) is related to a diminished CD69 and TGFbeta expression. PLoS One. 2009 Jun 22;4(6):e5981.

22. Varga J, Pasche B. Transforming growth factor beta as a therapeutic target in systemic sclerosis. Nat Rev Rheumatol. 2009 Apr;5(4):200-6.

23. LeRoy EC, Black C, Fleischmajer R, Jablonska S, Krieg T, Medsger TA Jr, Rowell N, Wollheim F. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. J Rheumatol. 1988 Feb;15(2):202-5.

24. Mok MY, Lau CS. The burden and measurement of cardiovascular disease in SSc. Nat Rev Rheumatol. 2010 Jul; 6(7):430-4. Epub 2010 Jun 1.

25. Ferri C, Valentini G, Cozzi F, et al. Systemic sclerosis: demographic, clinical, and serologic features and survival in 1,012 Italian patients. Medicine (Baltimore) 2002; 81:139.

26. Scussel-Lonzetti L, Joyal F, Raynauld JP, et al. Predicting mortality in systemic sclerosis: analysis of a cohort of 309 French Canadian patients with emphasis on features at diagnosis as predictive factors for survival. Medicine (Baltimore) 2002; 81:154.

27. Steen V, Medsger TA Jr. Predictors of isolated pulmonary hypertension in patients with systemic sclerosis and limited cutaneous involvement. Arthritis Rheum 2003; 48:516.

28. Trad S, Amoura Z, Beigelman C, et al. Pulmonary arterial hypertension is a major mortality factor in diffuse systemic sclerosis, independent of interstitial lung disease. Arthritis Rheum 2006; 54:184.

29. Nietert PJ, Silver RM. Systemic sclerosis: environmental and occupational risk factors. Curr Opin Rheumatol. 2000 Nov;12(6):520-6.

30. McCormic ZD, Khuder SS, Aryal BK, Ames AL, Khuder SA. Occupational silica exposure as a risk factor for scleroderma: a meta-analysis. Int Arch Occup Environ Health. 2010 Oct;83(7): 763-9. Epub 2010 Jan 3.

31. Englert H, Small-McMahon J, Chambers P, O'Connor H, Davis K, Manolios N, White R, Dracos G, Brooks P. Familial risk estimation in systemic sclerosis. Aust N Z J Med. 1999 Feb;29(1):36-41.

32. Arnett FC, Cho M, Chatterjee S, Aguilar MB, Reveille JD, Mayes MD. Familial occurrence frequencies and relative risks for systemic sclerosis (scleroderma) in three United States cohorts. Arthritis Rheum. 2001 Jun;44(6):1359-62.

33. Hudson M, Rojas-Villarraga A, Coral-Alvarado P, López-Guzmán S, Mantilla RD, Chalem P; Canadian Scleroderma Research Group; Colombian Scleroderma Research Group, Baron M, Anaya JM. Polyautoimmunity and familial autoimmunity in systemic sclerosis. J Autoimmun. 2008 Sep;31(2):156-9. Epub 2008 Jul 21.

34. Zorina VIa, Zorin SP. Scleroderma in 5-year-old twin girls. Pediatriia. 1982 Jul; (7):69-70.

35. De Keyser F, Peene I, Joos R, Naeyaert JM, Messiaen L, Veys EM. Occurrence of scleroderma in monozygotic twins. J Rheumatol. 2000 Sep;27(9):2267-9.

36. Feghali-Bostwick C, Medsger TA Jr, Wright TM. Analysis of systemic sclerosis in twins reveals low concordance for disease and high concordance for the presence of antinuclear antibodies. Arthritis Rheum. 2003 Jul;48(7):1956-63.

37. Kallenberg CG, Wouda AA, Hoet MH, van Venrooij WJ. Development of connective tissue disease in patients presenting with Raynaud's phenomenon: a six year follow up with emphasis on the predictive value of antinuclear antibodies as detected by immunoblotting. Ann Rheum Dis. 1988 Aug;47(8):634-41.

38. Zhou X, Tan FK, Xiong M, Arnett FC, Feghali-Bostwick CA. Monozygotic twins

clinically discordant for scleroderma show concordance for fibroblast gene expression profiles. Arthritis Rheum. 2005 Oct;52(10):3305-14.

39. Emerit I, Housset E, Feingold J. Chromosomal breakage and scleroderma: studies in family members. J Lab Clin Med. 1976 Jul;88(1):81-6.

40. Emerit I, Levy A, Housset E. Breakage factor in systemic sclerosis and protector effect of L-cysteine. Humangenetik. 1974;25(3):221-6.

41. Haaf T, Sumner AT, Köhler J, Willard HF, Schmid M. A microchromosome derived from chromosome 11 in a patient with the CREST syndrome of scleroderma. Cytogenet Cell Genet. 1992;60(1):12-7.

42. Sherer GK, Jackson BB, Leroy EC. Chromosome breakage and sister chromatid exchange frequencies in scleroderma. Arthritis Rheum. 1981 Nov; 24(11):1409-13.

43. Artlett CM, Black CM, Briggs DC, Stevens CO, Welsh KI. Telomere reduction in scleroderma patients: a possible cause for chromosomal instability. Br J Rheumatol. 1996 Aug; 35(8):732-7.

44. MacIntyre A, Brouilette SW, Lamb K, Radhakrishnan K, McGlynn L, Chee MM, Parkinson EK, Freeman D, Madhok R, Shiels PG. Association of increased telomere lengths in limited scleroderma, with a lack of age-related telomere erosion. Ann Rheum Dis. 2008 Dec; 67(12):1780-2. Epub 2008 Jul 28.

45. Tarhan F, Vural F, Kosova B, Aksu K, Cogulu O, Keser G, Gündüz C, Tombuloglu M, Oder G, Karaca E, Doganavsargil E. Telomerase activity in connective tissue diseases: elevated in rheumatoid arthritis, but markedly decreased in systemic sclerosis. Rheumatol Int. 2008 Apr;28(6):579-83. Epub 2007 Oct 16.

46. Graves JA, Disteche CM, Toder R. Gene dosage in the evolution and function of mammalian sex chromosomes. Cytogenet Cell Genet. 1998;80(1-4): 94-103.

47. LYON MF. Gene action in the X-chromosome of the mouse (Mus musculus L.). Nature. 1961 Apr 22;190:372-3.

48. Kristiansen M, Knudsen GP, Bathum L, Naumova AK, Sørensen TI, Brix TH, Svendsen AJ, Christensen K, Kyvik KO, Ørstavik KH. Twin study of genetic and aging effects on X chromosome inactivation. Eur J Hum Genet. 2005 May;13(5):599-606.

49. Ozbalkan Z, Bagişlar S, Kiraz S, Akyerli CB, Ozer HT, Yavuz S, Birlik AM, Calgüneri M, Ozçelik T. Skewed X chromosome inactivation in blood cells of women with scleroderma. Arthritis Rheum. 2005 May;52(5):1564-70.

50. Uz E, Loubiere LS, Gadi VK, Ozbalkan Z, Stewart J, Nelson JL, Ozcelik T. Skewed X-chromosome inactivation in scleroderma. Clin Rev Allergy Immunol. 2008 Jun;34(3):352-5.

51. Broen JC, Wolvers-Tettero IL, Geurts-van Bon L, Vonk MC, Coenen MJ, Lafyatis R, Radstake TR, Langerak AW. Skewed X chromosomal inactivation impacts T regulatory cell function in systemic sclerosis. Ann Rheum Dis. 2010 Dec; 69(12):2213-6. Epub 2010 Aug 10.

52. Invernizzi P, Miozzo M, Selmi C, Persani L, Battezzati PM, Zuin M, Lucchi S, Meroni PL, Marasini B, Zeni S, Watnik M, Grati FR, Simoni G, Gershwin ME, Podda M.X chromosome monosomy: a common mechanism for autoimmune diseases. J Immunol. 2005 Jul 1;175(1): 575-8.

53. Arnett FC, Bias WB, McLean RH, Engel M, Duvic M, Goldstein R, Freni-Titulaer L, McKinley TW, Hochberg MC.Connective tissue disease in southeast Georgia. A community based study of immunogenetic markers and autoantibodies. J Rheumatol. 1990 Aug; 17(8):1029-35.

54. Gladman DD, Keystone EC, Baron M, Lee P, Cane D, Mervert H. Increased frequency of HLA-DR5 in scleroderma. Arthritis Rheum. 1981 Jun;24(6):854-6.

55. Kuwana M, Kaburaki J, Okano Y, Inoko H, Tsuji K.The HLA-DR and DQ genes control the autoimmune response to DNA topoisomerase I in systemic sclerosis (scleroderma). J Clin Invest. 1993 Sep;92(3):1296-301.

56. Tan FK, Stivers DN, Arnett FC, Chakraborty R, Howard R, Reveille JD. HLA haplotypes and microsatellite polymorphisms in and around the major histocompatibility complex region in a Native American population with a high prevalence of scleroderma (systemic sclerosis). Tissue Antigens. 1999 Jan; 53(1):74-80.

57. Vlachoyiannopoulos PG, Dafni UG, Pakas I, Spyropoulou-Vlachou M, Stavropoulos-Giokas C, Moutsopoulos HM. Systemic scleroderma in Greece: low mortality and strong linkage with HLA-DRB1*1104 allele. Ann Rheum Dis. 2000 May;59(5):359-67.

58. Reveille JD, Durban E, MacLeod-St Clair MJ, Goldstein R, Moreda R, Altman RD, Arnett FC. Association of amino acid sequences in the HLA-DQB1 first domain with antitopoisomerase I autoantibody response in scleroderma (progressive systemic sclerosis). J Clin Invest. 1992 Sep;90(3):973-80.

59. Arnett FC, Gourh P, Shete S, Ahn CW, Honey RE, Agarwal SK, Tan FK, McNearney T, Fischbach M, Fritzler MJ, Mayes MD, Reveille JD. Major histocompatibility complex (MHC) class II alleles, haplotypes and epitopes which confer susceptibility or protection in systemic sclerosis: analyses in 1300 Caucasian, African-American and Hispanic cases and 1000 controls. Ann Rheum Dis. 2010 May;69(5):822-7. Epub 2009 Jul 12.

60. Simeón CP, Fonollosa V, Tolosa C, Palou E, Selva A, Solans R, Armadans L, Moreno E, Marsal S, Vilardell M. Association of HLA class II genes with systemic sclerosis in Spanish patients. J Rheumatol. 2009 Dec;36(12):2733-6. Epub 2009 Nov 2.

61. Karp DR, Marthandan N, Marsh SG, Ahn C, Arnett FC, Deluca DS, Diehl AD, Dunivin R, Eilbeck K, Feolo M, Guidry PA, Helmberg W, Lewis S, Mayes MD, Mungall C, Natale DA, Peters B, Petersdorf E, Reveille JD, Smith B, Thomson G, Waller MJ, Scheuermann RH. Novel sequence feature variant type analysis of the HLA genetic association in systemic sclerosis. Hum Mol Genet. 2010 Feb 15;19(4):707-19. Epub 2009 Nov 18.

62. Beretta L, Santaniello A, Mayo M, Cappiello F, Marchini M, Scorza R. A 3-factor epistatic model predicts digital ulcers in Italian scleroderma patients. Eur J Intern Med. 2010 Aug;21(4): 347-53. Epub 2010 Jun 23.

63. Nguyen B, Mayes MD, Arnett FC, del Junco D, Reveille JD, Gonzalez EB, Draeger HT, Perry M, Hendiani A, Anand KK, Assassi S. HLA-DRB1*0407 and *1304 are risk factors for scleroderma renal crisis. Arthritis Rheum. 2011 Feb; 63(2):530-4. doi: 10.1002/art.30111.

64. Remmers EF, Plenge RM, Lee AT, Graham RR, Hom G, Behrens TW, de Bakker PI, Le JM, Lee HS, Batliwalla F, Li W, Masters SL, Booty MG, Carulli JP, Padyukov L, Alfredsson L, Klareskog L, Chen WV, Amos CI, Criswell LA, Seldin MF, Kastner DL, Gregersen PK. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N Engl J Med. 2007 Sep 6;357(10):977-86.

65. Rueda B, Broen J, Simeon C, Hesselstrand R, Diaz B, Suárez H, Ortego-Centeno N, Riemekasten G, Fonollosa V, Vonk MC, van den Hoogen FH, Sanchez-Román J, Aguirre-Zamorano MA, García-Portales R, Pros A, Camps MT, Gonzalez-Gay MA, Coenen MJ, Airo P, Beretta L, Scorza R, van Laar J, Gonzalez-Escribano MF, Nelson JL, Radstake TR, Martin J. The STAT4 gene influences the genetic predisposition to systemic sclerosis phenotype. Hum Mol Genet. 2009 Jun 1;18(11):2071-7. Epub 2009 Mar 13.

66. Tsuchiya N, Kawasaki A, Hasegawa M, Fujimoto M, Takehara K, Kawaguchi Y, Kawamoto M, Hara M, Sato S. Association of STAT4 polymorphism with systemic sclerosis in a Japanese population. Ann Rheum Dis. 2009 Aug; 68(8):1375-6.

67. Dieudé P, Guedj M, Wipff J, Ruiz B, Hachulla E, Diot E, Granel B, Sibilia J, Tiev K, Mouthon L, Cracowski JL, Carpentier PH, Amoura Z, Fajardy I, Avouac J, Meyer O, Kahan A, Boileau C, Allanore Y. STAT4 is a genetic risk factor for systemic sclerosis having additive effects with IRF5 on disease susceptibility and related pulmonary fibrosis. Arthritis Rheum. 2009 Aug;60(8): 2472-9.

68. Gourh P, Agarwal SK, Divecha D, Assassi S, Paz G, Arora-Singh RK,

Reveille JD, Shete S, Mayes MD, Arnett FC, Tan FK. Polymorphisms in TBX21 and STAT4 increase the risk of systemic sclerosis: evidence of possible gene-gene interaction and alterations in Th1/ Th2 cytokines. Arthritis Rheum. 2009 Dec;60(12):3794-806.

69. Korman BD, Kastner DL, Gregersen PK, Remmers EF. STAT4: genetics, mechanisms, and implications for autoimmunity. Curr Allergy Asthma Rep. 2008 Sep;8(5):398-403.

70. Avouac J, Fürnrohr BG, Tomcik M, Palumbo K, Zerr P, Horn A, Dees C, Akhmetshina A, Beyer C, Distler O, Schett G, Allanore Y, Distler JH. Inactivation of the transcription factor STAT4 prevents inflammation-driven fibrosis in systemic sclerosis animal models. Arthritis Rheum. 2010 Nov 30. [Epub ahead of print]

71. Sigurdsson S, Nordmark G, Göring HH, Lindroos K, Wiman AC, Sturfelt G, Jönsen A, Rantapää-Dahlqvist S, Möller B, Kere J, Koskenmies S, Widén E, Eloranta ML, Julkunen H, Kristjansdottir H, Steinsson K, Alm G, Rönnblom L, Syvänen AC. Polymorphisms in the tyrosine kinase 2 and interferon regulatory factor 5 genes are associated with systemic lupus erythematosus. Am J Hum Genet. 2005 Mar;76(3):528-37. Epub 2005 Jan 18.

72. Dieudé P, Guedj M, Wipff J, Avouac J, Fajardy I, Diot E, Granel B, Sibilia J, Cabane J, Mouthon L, Cracowski JL, Carpentier PH, Hachulla E, Meyer O, Kahan A, Boileau C, Allanore Y. Association between the IRF5 rs2004640 functional polymorphism and systemic sclerosis: a new perspective for pulmonary fibrosis. Arthritis Rheum. 2009 Jan;60(1):225-33.

73. Ito I, Kawaguchi Y, Kawasaki A, Hasegawa M, Ohashi J, Hikami K, Kawamoto M, Fujimoto M, Takehara K, Sato S, Hara M, Tsuchiya N. Association of a functional polymorphism in the IRF5 region with systemic sclerosis in a Japanese population. Arthritis Rheum. 2009 Jun;60(6):1845-50.

74. Dieude P, Dawidowicz K, Guedj M, Legrain Y, Wipff J, Hachulla E, Diot E, Sibilia J, Mouthon L, Cabane J, Amoura Z, Crakowski JL, Carpentier P, Avouac J, Meyer O, Kahan A, Boileau C, Allanore Y. Phenotype-haplotype correlation of IRF5 in systemic sclerosis: role of 2 haplotypes in disease severity. J Rheumatol. 2010 May;37(5):987-92. Epub 2010 Mar 15.

75. Barnes BJ, Moore PA, Pitha PM. Virus-specific activation of a novel interferon regulatory factor, IRF-5, results in the induction of distinct interferon alpha genes. J Biol Chem. 2001 Jun 29;276(26):23382-90. Epub 2001 Apr 12.

76. Eloranta ML, Franck-Larsson K, Lövgren T, Kalamajski S, Rönnblom A, Rubin K, Alm GV, Rönnblom L. Type I interferon system activation and association with disease manifestations in systemic sclerosis. Ann Rheum Dis. 2010 Jul; 69(7):1396-402. Epub 2010 May 14.

77. Lakos G, Melichian D, Wu M, Varga J. Increased bleomycin-induced skin fibrosis in mice lacking the Th1-specific transcription factor T-bet. Pathobiology. 2006;73(5):224-37.

78. Aliprantis AO, Wang J, Fathman JW, Lemaire R, Dorfman DM, Lafyatis R, Glimcher LH. Transcription factor T-bet regulates skin sclerosis through its function in innate immunity and via IL-13. Proc Natl Acad Sci U S A. 2007 Feb 20;104(8):2827-30. Epub 2007 Feb 16.

79. Gourh P, Agarwal SK, Divecha D, Assassi S, Paz G, Arora-Singh RK, Reveille JD, Shete S, Mayes MD, Arnett FC, Tan FK. Polymorphisms in TBX21 and STAT4 increase the risk of systemic sclerosis: evidence of possible gene-gene interaction and alterations in Th1/ Th2 cytokines. Arthritis Rheum. 2009 Dec;60(12):3794-806.

80. Yokoyama K, Su Ih IH, Tezuka T, Yasuda T, Mikoshiba K, Tarakhovsky A, Yamamoto T. BANK regulates BCR-induced calcium mobilization by promoting tyrosine phosphorylation of IP(3) receptor. EMBO J. 2002 Jan 15;21(1-2):83-92.

81. Orozco G, Abelson AK, González-Gay MA, Balsa A, Pascual-Salcedo D, García A, Fernández-Gutierrez B, Petersson I, Pons-Estel B, Eimon A, Paira S, Scherbarth HR, Alarcón-Riquelme M, Martín J. Study of functional variants of the BANK1 gene in rheumatoid arthritis. Arthritis Rheum. 2009 Feb;60(2):372-9.

82. Kozyrev SV, Abelson AK, Wojcik J, Zaghlool A, Linga Reddy MV, Sanchez E, Gunnarsson I, Svenungsson E, Sturfelt G, Jönsen A, Truedsson L, Pons-Estel BA, Witte T, D'Alfonso S, Barizzone N, Danieli MG, Gutierrez C, Suarez A, Junker P, Laustrup H, González-Escribano MF, Martin J, Abderrahim H, Alarcón-Riquelme ME. Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus. Nat Genet. 2008 Feb; 40(2):211-6. Epub 2008 Jan 20.

83. Rueda B, Gourh P, Broen J, Agarwal SK, Simeon C, Ortego-Centeno N, Vonk MC, Coenen M, Riemekasten G, Hunzelmann N, Hesselstrand R, Tan FK, Reveille JD, Assassi S, Garcia-Hernandez FJ, Carreira P, Camps M, Fernandez-Nebro A, Garcia de la Peña P, Nearney T, Hilda D, Gónzalez-Gay MA, Airo P, Beretta L, Scorza R, Radstake TR, Mayes MD, Arnett FC, Martin J. BANK1 functional variants are associated with susceptibility to diffuse systemic sclerosis in Caucasians. Ann Rheum Dis. 2010 Apr; 69(4):700-5. Epub 2009 Oct 8.

84. Dieudé P, Wipff J, Guedj M, Ruiz B, Melchers I, Hachulla E, Riemekasten G, Diot E, Hunzelmann N, Sibilia J, Tiev K, Mouthon L, Cracowski JL, Carpentier PH, Distler J, Amoura Z, Tarner I, Avouac J, Meyer O, Kahan A, Boileau C, Allanore Y. BANK1 is a genetic risk factor for diffuse cutaneous systemic sclerosis and has additive effects with IRF5 and STAT4. Arthritis Rheum. 2009 Nov; 60(11):3447-54.

85. Dymecki SM, Niederhuber JE, Desiderio SV. Specific expression of a tyrosine kinase gene, blk, in B lymphoid cells. Science. 1990 Jan 19;247(4940):332-6.

86. Gourh P, Agarwal SK, Martin E, Divecha D, Rueda B, Bunting H, Assassi S, Paz G, Shete S, McNearney T, Draeger H, Reveille JD, Radstake TR, Simeon CP, Rodriguez L, Vicente E, Gonzalez-Gay MA, Mayes MD, Tan FK, Martin J, Arnett FC. Association of the C8orf13-BLK region with systemic sclerosis in North-American and European populations. J Autoimmun. 2010 Mar;34(2):155-62. Epub 2009 Sep 30.

87. Ito I, Kawaguchi Y, Kawasaki A, Hasegawa M, Ohashi J, Kawamoto M, Fujimoto M, Takehara K, Sato S, Hara M, Tsuchiya N. Association of the FAM167A-BLK region with systemic sclerosis. Arthritis Rheum. 2010 Mar;62(3):890-5.

88. Redmond WL, Ruby CE, Weinberg AD. The role of OX40-mediated co-stimulation in T-cell activation and survival. Crit Rev Immunol. 2009;29(3): 187-201.

89. Gourh P, Arnett FC, Tan FK, Assassi S, Divecha D, Paz G, McNearney T, Draeger H, Reveille JD, Mayes MD, Agarwal SK. Association of TNFSF4 (OX40L) polymorphisms with susceptibility to systemic sclerosis. Ann Rheum Dis. 2010 Mar;69(3):550-5. Epub 2009 Sep 23.

90. Bossini-Castillo L, Broen JC, Simeon CP, Beretta L, Vonk MC, Ortego-Centeno N, Espinosa G, Carreira P, Camps MT, Navarrete N, González-Escribano MF, Vicente-Rabaneda E, Rodríguez L, Tolosa C, Román-Ivorra JA, Gómez-Gracia I, García-Hernández FJ, Castellví I, Gallego M, Fernández-Nebro A, García-Portales R, Egurbide MV, Fonollosa V, de la Peña PG, Pros A, González-Gay MA, Hesselstrand R, Riemekasten G, Witte T, Coenen MJ, Koeleman BP, Houssiau F, Smith V, de Keyser F, Westhovens R, De Langhe E, Voskuyl AE, Schuerwegh AJ, Chee MM, Madhok R, Shiels P, Fonseca C, Denton C, Claes K, Padykov L, Nordin A, Palm O, Lie BA, Airó P, Scorza R, van Laar JM, Hunzelmann N, Kreuter A, Herrick A, Worthington J, Radstake TR, Martín J, Rueda B. A replication study confirms the association of TNFSF4 (OX40L) polymorphisms with systemic sclerosis in a large European cohort. Ann Rheum Dis. 2010 Dec 27. [Epub ahead of print]

91. Radstake TR, van Bon L, Broen J, Hussiani A, Hesselstrand R, Wuttge DM, Deng Y, Simms R, Lubberts E, Lafyatis R. The pronounced Th17 profile in systemic sclerosis (SSc) together with intracellular expression of TGFbeta and IFNgamma distinguishes SSc phenotypes. PLoS One. 2009 Jun 17;4(6):e5903.

92. Faragó B, Magyari L, Sáfrány E, Csöngei V, Járomi L, Horvatovich K, Sipeky C, Maász A, Radics J, Gyetvai A, Szekanecz Z, Czirják L, Melegh B. Functional variants of interleukin-23 receptor gene confer risk for rheumatoid arthritis but not for systemic sclerosis. Ann Rheum Dis. 2008 Feb;67(2):248-50.

Epub 2007 Jul 2.

93. Rueda B, Broen J, Torres O, Simeon C, Ortego-Centeno N, Schijvenaars MM, Vonk MC, Fonollosa V, van den Hoogen FH, Coenen MJ, Sanchez-Román J, Aguirre-Zamorano MA, García-Portales R, Pros A, Camps MT, Gonzalez-Gay MA, Martin J, Radstake TR. The interleukin 23 receptor gene does not confer risk to systemic sclerosis and is not associated with systemic sclerosis disease phenotype. Ann Rheum Dis. 2009 Feb;68(2):253-6. Epub 2008 Aug 19.

94. Agarwal SK, Gourh P, Shete S, Paz G, Divecha D, Reveille JD, Assassi S, Tan FK, Mayes MD, Arnett FC. Association of interleukin 23 receptor polymorphisms with anti-topoisomerase-I positivity and pulmonary hypertension in systemic sclerosis. J Rheumatol. 2009 Dec; 36(12):2715-23. Epub 2009 Nov 16.

95. Talal N. Oncogenes, autogenes, and rheumatic diseases. Arthritis Rheum. 1994 Oct;37(10):1421-2.

96. Wetzig T, Petri JB, Mittag M, Haustein UF. Serum levels of soluble Fas/APO-1 receptor are increased in systemic sclerosis. Arch Dermatol Res. 1998 Apr; 290(4):187-90.

97. Broen J, Gourh P, Rueda B, Coenen M, Mayes M, Martin J, Arnett FC, Radstake TR; European Consortium on Systemic Sclerosis Genetics. The FAS -670A>G polymorphism influences susceptibility to systemic sclerosis phenotypes. Arthritis Rheum. 2009 Dec;60(12):3815-20.

98. Liakouli V, Manetti M, Pacini A, Tolusso B, Fatini C, Toscano A, Cipriani P, Guiducci S, Bazzichi L, Codullo V, Ruocco L, Dell'orso L, Carubbi F, Marrelli A, Abbate R, Bombardieri S, Ferraccioli G, Montecucco C, Valentini G, Matucci-Cerinic M, Ibba-Manneschi L, Giacomelli R. The -670G>A polymorphism in the FAS gene promoter region influences the susceptibility to systemic sclerosis. Ann Rheum Dis. 2009 Apr;68(4):584-90. Epub 2008 Apr 29.

99. Leask A, Denton CP, Abraham DJ. Insights into the molecular mechanism of chronic fibrosis: the role of connective tissue growth factor in scleroderma. J Invest Dermatol. 2004 Jan;122(1):1-6.

100. Fonseca C, Lindahl GE, Ponticos M, Sestini P, Renzoni EA, Holmes AM, Spagnolo P, Pantelidis P, Leoni P, McHugh N, Stock CJ, Shi-Wen X, Denton CP, Black CM, Welsh KI, du Bois RM, Abraham DJ. A polymorphism in the CTGF promoter region associated with systemic sclerosis. N Engl J Med. 2007 Sep 20;357(12):1210-20.

101. Gourh P, Mayes MD, Arnett FC. CTGF polymorphism associated with systemic sclerosis. N Engl J Med. 2008 Jan 17;358(3):308-9; author reply 309.

102. Rueda B, Simeon C, Hesselstrand R, Herrick A, Worthington J, Ortego-Centeno N, Riemekasten G, Fonollosa V, Vonk MC, van den Hoogen FH, Sanchez-Román J, Aguirre-Zamorano MA, García-Portales R, Pros A, Camps MT, Gonzalez-Gay MA, Gonzalez-Escribano MF, Coenen MJ, Lambert N, Nelson JL, Radstake TR, Martin J. A large multicentre analysis of CTGF -945 promoter polymorphism does not confirm association with systemic sclerosis susceptibility or phenotype. Ann Rheum Dis. 2009 Oct;68(10):1618-20. Epub 2008 Dec 3.

103. Kawaguchi Y, Ota Y, Kawamoto M, Ito I, Tsuchiya N, Sugiura T, Katsumata Y, Soejima M, Sato S, Hasegawa M, Fujimoto M, Takehara K, Kuwana M, Yamanaka H, Hara M. Association study of a polymorphism of the CTGF gene and susceptibility to systemic sclerosis in the Japanese population. Ann Rheum Dis. 2009 Dec;68(12):1921-4. Epub 2008 Dec 3.

104. Granel B, Argiro L, Hachulla E, Fajardy I, Weiller PJ, Durand JM, Frances Y, Dombey AM, Marquet S, Lesavre N, Disdier P, Bernard F, Hatron PY, Chevillard C. Association between a CTGF gene polymorphism and systemic sclerosis in a French population. J Rheumatol. 2010 Feb;37(2):351-8. Epub 2009 Dec 23.

105. Diaz-Gallo L, Gourh P, Broen J, Simeon C, Fonollosa V, Ortego-Centeno N, Agarwal S, Vonk M, Coenen M, Riemekasten G, Hunzelmann N, Hesselstrand R, Tan F, Reveille J, Assassi S, García-Hernandez F, Carreira P, Camps M, Fernandez-Nebro A, de la Peña PG, Nearney T, Hilda D, González-Gay M, Airo P, Beretta L, Scorza R,

Herrick A, Worthington J, Pros A, Gómez-Gracia I, Trapiella L, Espinosa G, Castellvi I, Witte T, de Keyser F, Vanthuyne M, Mayes M, Radstake T, Arnett F, Martin J, Rueda B. Analysis of the influence of PTPN22 gene polymorphisms in systemic sclerosis. Ann Rheum Dis. 2011 Mar;70(3):454-62. Epub 2010 Dec 3.

106. Dieudé P, Guedj M, Truchetet ME, Wipff J, Revillod L, Riemekasten G, Matucci-Cerinic M, Melchers I, Hachulla E, Airo P, Diot E, Hunzelmann N, Mouthon L, Cabane J, Cracowski JL, Riccieri V, Distler J, Amoura Z, Valentini G, Camaraschi P, Tarner I, Frances C, Carpentier P, Brembilla NC, Meyer O, Kahan A, Chizzolini C, Boileau C, Allanore Y. Association of the CD226 307Ser variant with systemic sclerosis: Evidence for a contribution of co-stimulation pathways in SSc pathogenesis. Arthritis Rheum. 2010 Dec 15. doi: 10.1002/art.30204. [Epub ahead of print]

107. Dieudé P, Guedj M, Wipff J, Ruiz B, Riemekasten G, Airo P, Melchers I, Hachulla E, Cerinic MM, Diot E, Hunzelmann N, Caramaschi P, Sibilia J, Tiev K, Mouthon L, Riccieri V, Cracowski JL, Carpentier PH, Distler J, Amoura Z, Tarner I, Avouac J, Meyer O, Kahan A, Boileau C, Allanore Y. NLRP1 influences the systemic sclerosis phenotype: a new clue for the contribution of innate immunity in systemic sclerosis-related fibrosing alveolitis pathogenesis. Ann Rheum Dis. 2010 Dec 13. [Epub ahead of print]

108. Manetti M, Allanore Y, Revillod L, Fatini C, Guiducci S, Cuomo G, Bonino C, Riccieri V, Bazzichi L, Liakouli V, Cipriani P, Giacomelli R, Abbate R, Bombardieri S, Valesini G, Montecucco C, Valentini G, Ibba-Manneschi L, Matucci-Cerinic M. A genetic variation located in the promoter region of the UPAR (CD87) gene is associated with the vascular complications of systemic sclerosis. Arthritis Rheum. 2011 Jan;63(1):247-56. doi: 10.1002/art.30101.

109. Zhou X, Tan FK, Wang N, Xiong M, Maghidman S, Reveille JD, Milewicz DM, Chakraborty R, Arnett FC. Genome-wide association study for regions of systemic sclerosis susceptibility in a Choctaw Indian population with high disease prevalence. Arthritis Rheum. 2003 Sep; 48(9):2585-92.

110. Zhou X, Lee JE, Arnett FC, Xiong M, Park MY, Yoo YK, Shin ES, Reveille JD, Mayes MD, Kim JH, Song R, Choi JY, Park JA, Lee YJ, Lee EY, Song YW, Lee EB. HLA-DPB1 and DPB2 are genetic loci for systemic sclerosis: a genome-wide association study in Koreans with replication in North Americans. Arthritis Rheum. 2009 Dec;60(12):3807-14.

111. Radstake TR, Gorlova O, Rueda B, Martin JE, Alizadeh BZ, Palomino-Morales R, Coenen MJ, Vonk MC, Voskuyl AE, Schuerwegh AJ, Broen JC, van Riel PL, van 't Slot R, Italiaander A, Ophoff RA, Riemekasten G, Hunzelmann N, Simeon CP, Ortego-Centeno N, González-Gay MA, González-Escribano MF; Spanish Scleroderma Group, Airo P, van Laar J, Herrick A, Worthington J, Hesselstrand R, Smith V, de Keyser F, Houssiau F, Chee MM, Madhok R, Shiels P, Westhovens R, Kreuter A, Kiener H, de Baere E, Witte T, Padykov L, Klareskog L, Beretta L, Scorza R, Lie BA, Hoffmann-Vold AM, Carreira P, Varga J, Hinchcliff M, Gregersen PK, Lee AT, Ying J, Han Y, Weng SF, Amos CI, Wigley FM, Hummers L, Nelson JL, Agarwal SK, Assassi S, Gourh P, Tan FK, Koeleman BP, Arnett FC, Martin J, Mayes MD. Genome-wide association study of systemic sclerosis identifies CD247 as a new susceptibility locus. Nat Genet. 2010 May;42(5):426-9.

112. Gorman CL, Russell AI, Zhang Z, Cunninghame Graham D, Cope AP, Vyse TJ. Polymorphisms in the CD3Z gene influence TCRzeta expression in systemic lupus erythematosus patients and healthy controls. J Immunol. 2008 Jan 15;180(2):1060-70.

113. Dieudé P, Boileau C, Guedj M, Avouac J, Ruiz B, Hachulla E, Diot E, Cracowski JL, Tiev K, Sibilia J, Mouthon L, Frances C, Amoura Z, Carpentier P, Cosnes A, Meyer O, Kahan A, Chiocchia G, Allanore Y. Independent replication establishes the CD247 gene as a genetic systemic sclerosis susceptibility factor. Ann Rheum Dis. 2011 Apr 7. [Epub ahead of print]

114. Gorlova O, Martin JE, Rueda B, Koeleman BPC, Ying J, Teruel M, Diaz-Gallo LM, Broen J, Vonk MC, Simeon CP, Alizadeh BZ, Coenen MJH, Voskuyl AE, Schuerwegh AJ, Van Riel PLCM,

Vanthuyne M, Van ‘t Slot R, Italiaander, Ophoff RA, Hunzelmann N, Fonol- losa V, Ortego-Centeno N, González-Gay MA, García-Hernández FJ, González-Escribano MF, Airo’ P, Van Laar J, Worthington J, Hesselstrand R, Smith V, De Keyser F, Houssiau F, Chee MM, Madhok R, Shiels P, Westhovens R, Kreuter A, De Baere E, Witte T, Padyukov L, Nordin A, Scorza R, Lunardi C, Lie BA, Hoffmann-Vold AM, De la Peña PG, Carreira P, Varga J, Hinchcliff M, Lee AT, Gourh P, Amos CI, Wigley FM, Hummers LK, Nelson JL, Riemekasten G, Herrick A, Beretta L, Fonseca C, Denton CP, Gregersen PK, Agarwal S, Assassi S, Tan FK, Arnett FC, Radstake TRDJ, Mayes MD, Martin J. Identification of novel genetic markers associated with clinical phenotypes and autoantibody subsets of systemic sclerosis through a genome wide association strategy. PLoS Genet. 2011

115. Allanore Y, Saad M, Dieudé, Avouac J, Distler JHW, Amouye P, Matucci-Cerinic M, Riemekasten G, Airo’ P, Melchers I, Hachulla E, Daniele C, Wichmann HE, Wipff J, Lambert JC, Hunzelmann N, Tiev K, Caramaschi P, Diot E, Kowal-Bielecka O, Valentini G, Mouthon L, Czirják L, Damjanov N, Salvi E, Conti C, Muller M, Muller-Ladner U, Riccieri V, Ruiz B, Cracowski JL, Letenneur L, Dupuy AM , Meyer O, Kahan A, Munnich A, Boileau C, Martinez M. Genomewide scan identifies TNIP1, PSORS1C1 and RHOB as novel risk loci for Systemic Sclerosis. PLoS Genet. 2011

116. Broen JCA, Coenen MJH, Radstake TRDJ. Deciphering the genetic background of systemic sclerosis. Expert Rev Clin Immunol. 2011

117. Dieudé P, Guedj M, Wipff J, Ruiz B, Riemekasten G, Airo P, Melchers I, Hachulla E, Cerinic MM, Diot E, Hunzelmann N, Caramaschi P, Sibilia J, Tiev K, Mouthon L, Riccieri V, Cracowski JL, Carpentier PH, Distler J, Amoura Z, Tarner I, Avouac J, Meyer O, Kahan A, Boileau C, Allanore Y. NLRP1 influences the systemic sclerosis phenotype: a new clue for the contribution of innate immunity in systemic sclerosis-related fibrosing alveolitis pathogenesis. Ann Rheum Dis. 2010 Dec 13. [Epub ahead of print]

Part I

Genetic association studies

Paul Klee, *Intention* 1938

Chapter 2

The *STAT4* gene influences the genetic predisposition to systemic sclerosis phenotype.

Rueda B, Broen J, Simeon C, Hesselstrand R, Diaz B, Suárez H, Ortego-Centeno N, Riemekasten G, Fonollosa V, Vonk MC, van den Hoogen FH, Sanchez-Román J, Aguirre-Zamorano MA, García-Portales R, Pros A, Camps MT, Gonzalez-Gay MA, Coenen MJ, Airo' P, Beretta L, Scorza R, van Laar J, Gonzalez-Escribano MF, Nelson JL, Radstake TR, Martin J.

Human Molecular Genetics 2009

Abstract

The aim of this study was to investigate the possible role of *STAT4* gene in the genetic predisposition to systemic sclerosis (SSc) susceptibility or clinical phenotype. A total of 1317 SSc patients [896 with limited cutaneous SSc (lcSSc) and 421 with diffuse cutaneous SSc (dcSSc)] and 3113 healthy controls, from an initial case–control set of Spanish Caucasian ancestry and five independent cohorts of European ancestry (The Netherlands, Germany, Sweden, Italy and North-America), were included in the study. The rs7574865 polymorphism was selected as *STAT4* genetic marker. We observed that the rs7574865 T allele was significantly associated with susceptibility to lcSSc in the Spanish population (p = 1.9 × 10−5 odds ratio (OR) 1.61 95% confidence intervals (CI) 1.29–1.99), but not with dcSSc (p = 0.41 OR 0.84 95% CI 0.59–1.21). Additionally, a dosage effect was observed showing individuals with rs7574865 TT genotype higher risk for lcSSc (OR 3.34, p = 1.02 × 10−7 95% CI 2.11–5.31). The association of the rs7574865 T allele with lcSSc was confirmed in all the replication cohorts with different effect sizes (OR ranging between 1.15 and 1.86), as well as the lack of association of *STAT4* with dcSSc. A meta-analysis to test the overall effect of the rs7574865 polymorphism showed a strong risk effect of the T allele for lcSSc susceptibility (pooled OR 1.54 95% CI 1.36–1.74; P < 0.0001). Our data show a strong and reproducible association of the STAT4 gene with the genetic predisposition to lcSSc suggesting that this gene seems to be one of the genetic markers influencing SSc phenotype.

Introduction

Systemic sclerosis (SSc) is one of the most disabling autoimmune conditions characterized by an extensive fibrotic process that affects multiple organs and tissues (1). Although the etiology of SSc is still poorly understood, it is widely accepted that the interaction of environmental factors with different individual genetic factors leads to SSc development (2).

To date, our knowledge of genetic factors contributing to SSc susceptibility or clinical phenotypes is very limited. Only genes within the major histocompatibility complex have been associated with SSc genetic predisposition or clinical manifestations in a consistent and reproducible fashion (2). Similar to other complex genetic disorders, several loci are expected to contribute to SSc genetic predisposition each with only moderate magnitude (3).

Genes implicated in the main pathogenic mechanisms of SSc are interesting candidates. One of central events responsible of the SSc development is the dysregulation of the immune system. The altered immune response in SSc is marked by an increased T cell activation and the secretion of pro-inflammatory mediators that contribute to the generation of fibrosis and endothelial alterations, hallmarks of SSc (4).

Several mechanisms regulate T cell activation and differentiation, being one of the most important the specific activation of gene transcription after cytokine stimulation (5). Signal transducers and activators of transcription (STAT) are a family of transcription factors that have been demonstrated to exert a fundamental role in driving T cell differentiation and signaling (6,7). Among the six described STAT proteins, STAT4 has acquired great interest. STAT4 is implicated in the differentiation of Th1 and also in the recently described Th17 cells, two of the T cell subsets that are implicated in SSc pathogenesis (4,8–10). In addition, the *STAT4* gene has been associated with genetic predisposition to different autoimmune diseases (AIDs), such as rheumatoid arthritis (RA), systemic lupus erythematosus (SLE), type 1 diabetes, inflammatory bowel disease and Sjögren's syndrome. A tag

single nucleotide polymorphism (SNP) of *STAT4* intron 3 (rs7574865) showed the strongest effect (11–17).

In view of these findings, we considered the *STAT4* gene as an excellent candidate gene and aimed to investigate its possible implication in the genetic predisposition to SSc susceptibility or clinical phenotype.

Results

STAT4 is associated with limited cutaneous SSc in the Spanish population

First we conducted an association study in a case–control set of Spanish Caucasian ancestry. The distribution of the *STAT4* rs7574865 genotypes and alleles in the Spanish controls was very similar to that reported previously in Caucasian populations [minor allele frequency (MAF) 0.21] and were observed to be in Hardy–Weinberg equilibrium (HWE) (**Table 1**) (11). After comparing the genotypes and alleles of the rs7574865 polymorphism between SSc patients with limited disease and healthy controls, we observed that the rs7574865 T allele was strongly associated with lcSSc (p = 1.9 × 10–5 odds ratio (OR) 1.61 95% confidence intervals (CI) 1.29–1.99) (**Table 1)**. In addition, the TT genotype was significantly increased in lcSSc patients compared with controls (P = 1.02 × 10−7 OR 3.34 95% CI 2.11–5.31) (Table 1). Nevertheless, no evidence of association was observed when the distribution of *STAT4* rs7574865 genotypes or alleles between patients with diffuse disease and healthy controls was compared (Table 1).

The increased frequency of the rs7574865 T allele and genotype among lcSSc patients was observed also when they were compared with dcSSc patients (T allele in lcSSc 29.5% versus 23.3% in dcSSc), although the differences did not reach statistical significance (Table 1).

Table 1. Distribution of *STAT4* rs7574865 genetic variant in the Spanish population

rs7574865	Controls *n*= 1296	lcSSc *n*= 242	*P*	OR 95% CI	dcSSc *n*= 90	*P*	OR 95% CI
Genotype							
GG	813 (62.7)	130 (53.7)	0.008	0.68 (0.50–0.77)	54 (60.0)	0.61	0.88 (0.57–1.36)
GT	428 (33.0)	81 (33.5)	0.89	1.02 (0.76–1.37)	30 (33.3)	0.95	1.03 (0.65–1.62)
TT	55 (4.2)	31 (12.8)	1.02×10^{-7}	3.34 (2.11–5.31)	6 (6.7)	0.29	1.82 (0.78–4.24)
Allele							
G	2054 (79.2)	341 (70.5)	1.9×10^{-5}	0.62 (0.50–0.77)	138 (76.7)	0.41	0.84 (0.59–1.21)
T	538 (20.8)	143 (29.5)		1.61 (1.29–1.99)	42 (23.3)		1.18 (0.82–1.68)

A replication study in five independent Caucasian populations and a meta-analysis confirm that STAT4 is strongly associated with limited SSc

In view of the interesting findings observed in our Spanish population and in order to confirm the association of *STAT4* with lcSSc, we conducted a large replication study including five independent populations with Caucasian ancestry. All analyzed control populations were in HWE for the *STAT4* rs7574865 genetic variant. The frequency for the rs7574865 T allele in the control populations ranged between 0.21 and 0.25 as expected from previous studies (11).

Interestingly, we confirmed the significant association of the rs7574865 T allele with susceptibility to lcSSc in all cohorts tested including patients from the Netherlands (OR 1.8 95% CI 1.32–2.47), Germany (OR 1.6 95% CI 1.15–2.21), North America (OR 1.4 95% CI 1.04–3.66) and Italy (OR 2.0 95% CI 1.05–1.78) (Table 2). Although not statistical significant, we observed the same trend in the Swedish population (rs7574865 T allele OR 1.3). Probably, this was due to the lower frequency for both rs7574865 TT genotype and T allele observed in the Swedish lcSSc patients compared with that observed in the other populations analyzed. Furthermore, the lack of association of STAT4 rs7574865 polymorphism with the diffuse cutaneous SSc subtype was confirmed in all five independent case–control sets (**Table 2**).

Table 2. Distribution of *STAT4* rs7574865 genetic variant in five replication cohorts

Population	rs7574865	Controls	lcSSc	*P*	dcSSc	*P*
Dutch		*n* = 893	*n* = 101		*n* = 30	
	GG	552 (61.8)	49 (48.5)	0.009	18 (60.0)	0.84
	GT	295 (33.0)	37 (36.6)	0.46	10 (33.3)	0.97
	TT	46 (5.2)	15 (14.9)	1.2×10^{-4}	2 (6.7)	0.71
	G	1399 (78.3)	135 (66.8)	2.2×10^{-4}	46 (76.7)	0.76
	T	387 (21.7)	67 (33.2)		14 (23.3)	
German		*n* = 227	*n* = 153		*n* = 117	
	GG	138 (60.8)	77 (50.3)	0.04	63 (53.8)	0.22
	GT	78 (34.4)	57 (37.3)	0.56	46 (39.3)	0.36
	TT	11 (4.8)	19 (12.4)	0.007	8 (6.8)	0.44
	G	354 (78.0)	211 (69.0)	0.005	134 (74.4)	0.19
	T	100 (22.0)	95 (31.0)		46 (25.6)	
North-American		*n* = 77	*n* = 30		*n* = 53	
	GG	44 (57.1)	9 (30.0)	0.012	23 (43.4)	0.12
	GT	27 (35.1)	18 (60.0)	0.019	25 (47.2)	0.16
	TT	6 (7.8)	3 (10.0)	0.71	5 (9.4)	0.74
	G	115 (74.7)	36 (60.0)	0.034	71 (67.0)	0.17
	T	39 (25.3)	24 (40.0)		35 (33.0)	
Italian		*n* = 362	*n* = 259		*n* = 92	
	GG	222 (61.3)	140 (54.1)	0.07	48 (52.2)	0.11
	GT	127 (35.1)	99 (38.2)	0.42	39 (42.4)	0.19
	TT	13 (3.6)	20 (7.7)	0.024	5 (5.4)	0.42
	G	571 (78.9)	379 (73.2)	0.019	135 (73.4)	0.11
	T	153 (21.1)	139 (26.8)		49 (26.6)	
Swedish		*n* = 285	*n* = 111		*n* = 39	
	GG	169 (59.3)	58 (52.3)	0.20	24 (61.5)	0.79
	GT	100 (35.1)	44 (39.6)	0.39	12 (30.8)	0.59
	TT	16 (5.6)	9 (8.1)	0.36	3 (7.7)	0.60
	G	438 (76.8)	160 (72.1)	0.16	60 (76.9)	0.99
	T	132 (23.2)	62 (27.9)		18 (23.1)	

Additionally, we conducted a meta-analysis to test the overall effect of the rs7574865 T allele on lcSSc susceptibility. The estimation of the homogeneity between the Spanish, Dutch, German, North American, Italian and Swedish populations did not reveal significant differences between them. In consequence, we performed a combined analysis considering the data from the six case–control

series using the Mantel–Haenszel test under fixed effects that again showed the strong risk effect of the *STAT4* rs7574865 T allele for lcSSc genetic predisposition (pooled OR 1.54 95% CI 1.36–1.74; $p < 0.0001$) (**Fig. 1A**). Furthermore, in the pooled analysis, individuals carrying the TT genotype showed an increased risk for lcSSc susceptibility (pooled OR of 2.55 95% CI 1.93–3.36, $p < 0.0001$). The same trend was observed for the comparison of lcSSc with dcSSc that showed a significant increased frequency of the *STAT4* rs7574865 T allele and TT genotype in the lcSSc patients group (pooled OR for the T allele 1.27 95% CI 1.04–1.54, $p = 0.019$ (Fig. 1B); pooled OR for the TT genotype 1.7 95% CI 1.10–2.60, $p = 0.02$).

In addition, we performed a pooled analysis to further investigate the overall effect of the rs7574865 T allele on dcSSc susceptibility. The meta-analysis reached a pooled OR of 1.23 95% CI 1.01–1.46 ($p = 0.03$) showing a trend for association of the rs7574865 T allele with dcSSc when compared with controls, but with a lower effect than in lcSSc. However, the pooled analysis for the TT genotype reached no statistically significant association with dcSSc (pooled OR 1.44 95% CI 0.93–2.26; $p = 0.12$).

STAT4 is not implicated in SSc clinical manifestations

The possible implication of *STAT4* in SSc clinical manifestations was further investigated. First, we analyzed the Spanish cohort individuals in whom data regarding selective auto-antibodies status (306 SSc patients for anti-Scl-70 and 294 for ACA) were available. Data related to pulmonary function were available in 281 SSc patients for pulmonary fibrosis, 233 for DLCO and 279 for forced vital capacity (FVC). We observed no significant association between the rs7574865 genetic variant and the presence of SSc specific antibodies (Scl70 or ACA) or pulmonary involvement (Table 3). This lack of association of *STAT4* and SSc clinical features was confirmed in the five replication cohorts (data not shown).

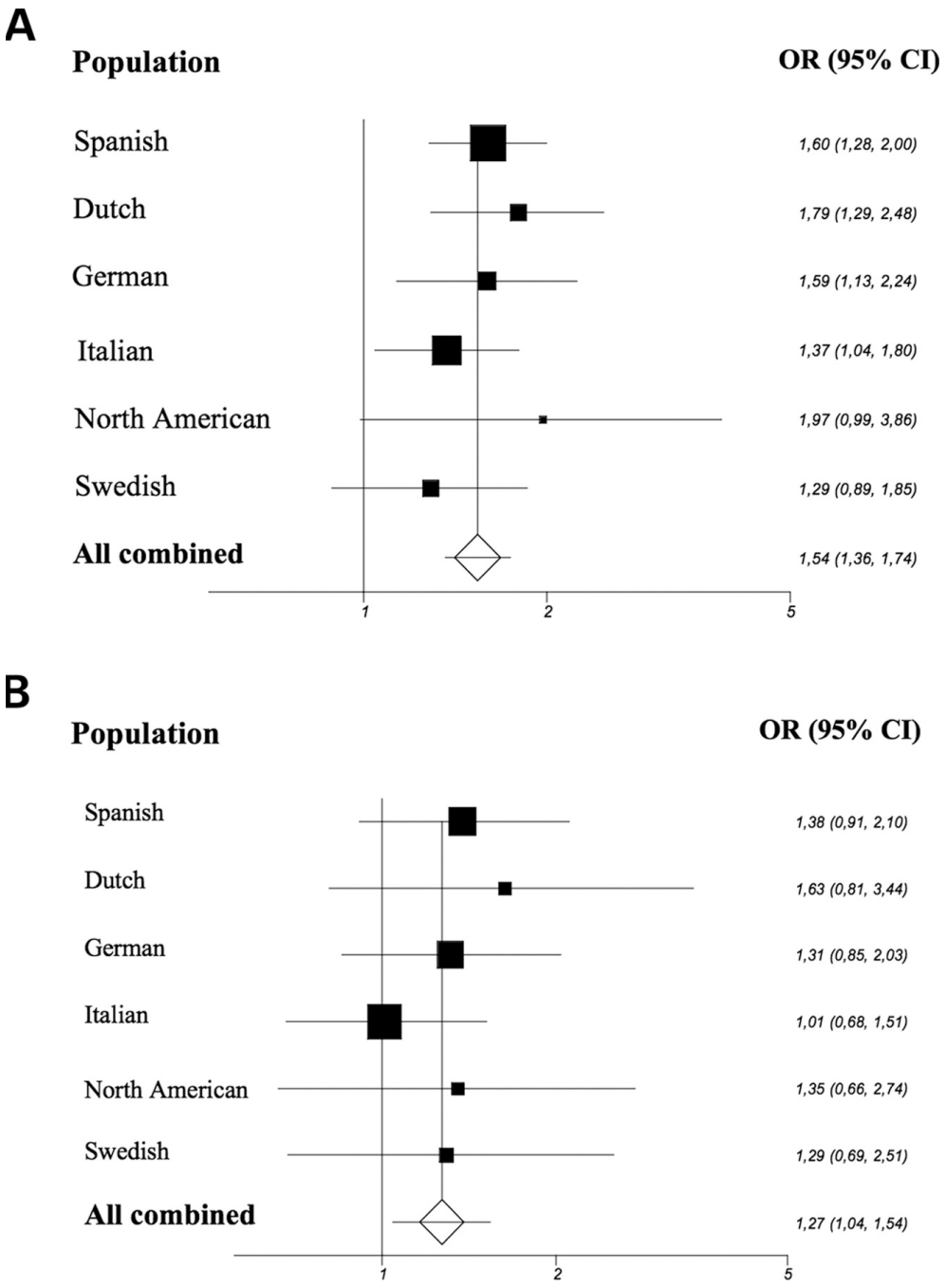


Figure 1. Effect of the *STAT4* rs7574865 T allele in lcSSc in Caucasian populations. (A) lcSSc versus controls. Mantel–Haenszel test for overall association of rs7574865 T allele with lcSSc susceptibility $p < 0.0001$. (B) lcSSc versus dcSSc. Mantel–Haenszel test for overall association of rs7574865 T allele in lcSSc susceptibility $p = 0.019$.

Table 3. Distribution of *STAT4* rs7574865 genetic variant according to SSc clinical features in the Spanish population

		Genotype			Allele	
rs7574865		**GG**	**GT**	**TT**	**G**	**T**
ACA+	n = 151 (%)	81 (53.6)	53 (35.1)	17 (11.3)	215 (71.2)	87 (28.8)
ACA−	n = 155 (%)	82 (52.9)	55 (35.5)	18 (11.6)	219 (70.6)	91 (29.4)
Anti-Scl70+	n = 55 (%)	30 (54.5)	21 (38.2)	4 (7.3)	81 (73.6)	29 (26.4)
Anti-Scl70−	n = 239 (%)	130 (54.4)	81 (33.9)	28 (11.7)	341 (71.3)	137 (28.7)
Pulmonary fibrosis+	n = 90 (%)	47 (52.2)	28 (31.1)	15 (16.7)	122 (67.8)	58 (32.2)
Pulmonary fibrosis−	n = 191 (%)	109 (57.1)	66 (34.6)	16 (8.4)	284 (74.3)	98 (25.7)
Predicted DLCO >75%	n = 129 (%)	74 (54.7)	39 (30.2)	16 (12.4)	187 (72.5)	71 (27.5)
Predicted DLCO <75%	n = 104 (%)	57 (54.8)	37 (35.6)	10 (9.6)	151 (72.6)	57 (27.4)
Predicted FVC >75%	n = 201 (%)	112 (55.7)	66 (32.8)	23 (11.4)	290 (72.1)	112 (27.9)
Predicted FVC <75%	n = 78 (%)	40 (51.3)	29 (37.2)	9 (11.5)	109 (69.9)	47 (30.1)

Discussion

SSc is characterized by skin fibrosis, vascular damage and immune system activation leading to a very heterogeneous pattern of clinical symptoms. Although, these three key events affect all SSc patients, there is a broad range of variability in the extent and severity of skin and internal organ involvement between individuals. In the clinical practice, SSc patients are divided in two groups characterized based upon the distribution of skin involvement (18).

One of the possible explanations for the existence of these two markedly different forms of disease could reside in the presence of differences in the genetic background between individuals. The results from the present study are in line with this hypothesis and support a role of the STAT4 gene as a genetic determinant of SSc phenotype. Through a large association study that included six independent populations of Caucasian ancestry, we demonstrate that the *STAT4* rs7574865 genetic variant confers susceptibility to lcSSc, but not to dcSSc. The rs7574865 T allele was significantly increased in the group of lcSSc patients compared with healthy controls or dcSSc patients. In addition, a dosage effect was observed showing that individuals with the rs7574865 TT genotype have a 2-fold increased risk for lcSSc susceptibility.

To our knowledge, this is the strongest genetic association with lcSSc described so far, being reproducible in different independent populations. This fact, together with the high statistical power of our study design, strongly suggests that the *STAT4* gene may be considered as a novel susceptibility gene for lcSSc.

Nevertheless, our data does not support evidence that the STAT4 gene exerts such a strong effect in dcSSc genetic predisposition. Certainly, the independent analyses of rs7574865 T allele and TT genotype and the meta-analysis of rs7574865 TT genotype in dcSSc compared with the healthy population point that this polymorphism has a negligible effect on dcSSc susceptibility. Only the meta-analysis of the rs7574865 T allele showed a trend for association with dcSSc. In this regard, considering that there is still no consensus on the levels of statistical significance for meta-analysis (19), the *p* value (0.03) and the 95% intervals close

to the null obtained for the meta-analysis of the rs7574865 T allele found in our study do not allow us to conclude an association of this SNP with dcSSc. In addition, the comparison of the overall distribution of rs7574865 T and TT genotype between lcSSc and dcSSc patients showed that this genetic marker confers a significant increased risk for SSc limited phenotype susceptibility. However, due to the lower proportion of dcSSc in European populations, the total number of dcSSc included in our study (n = 421) may not be enough to reach a high statistical power. On the other hand, it has been empirically demonstrated that meta-analyses with P-values between 0.01 and 0.05, similar to that obtained in our analyses for the dcSSc phenotype, do not have strong reliability (20). On this basis, the results from the present study suggest that the *STAT4* gene does not seem to play a major role in dcSS; however, further independent studies are needed to confirm this hypothesis.

STAT4 is an essential transcription factor for the regulation of the immune response. Upon the stimulation of cytokines such as IL-12, IL-23 or IL-17 STAT4 is activated and drives the expression of several pro-inflammatory mediators implicated in the differentiation and proliferation of Th1 and Th17 T cell subsets (8,21). Therefore, it is plausible that the elevated levels of IL-12, IL-23 and IL-17 observed in SSc patients lead to the activation of the STAT4 pathway (22,23). Then, a prolonged STAT4 increased activity due to different genetic variation in this gene might cause a sustained inflammatory response together with the expansion and infiltration of pro-inflammatory T cell subpopulations in skin and internal organs of SSc patients. However, further functional studies are necessary to elucidate the exact molecular mechanisms by which STAT4 is implicated in the pathogenesis of SSc and more precisely, how this transcription factor can lead to the development of the lcSSc.

In the past few years, using both genome-wide association studies and candidate gene association studies, genes such as *PTPN22*, *CTLA4* or *IL23R* have been associated with genetic susceptibility to various autoimmune conditions, leading to the hypothesis that different AIDs may share common genetic factors and pathways (24–27).

The association of *STAT4* with SSc reported here, together with previous findings that have shown recently the association of *STAT4* with susceptibility to RA, SLE and Sjögren syndrome support the notion that this gene seems to be another common genetic factor for autoimmunity. Nevertheless, to understand how *STAT4* influences the development of AIDs the next step is to conduct functional studies in order to identify which is/are the real causing variants that may influence *STAT4* activity or expression.

In summary, in this study we described for the first time a strong and reproducible association of the *STAT4* gene with the genetic susceptibility to lcSSc. Our data shed light on the pathogenic mechanisms that may underlie SSc development and open a new opportunity for the treatment of this debilitating disease.

Methods

Patients

A total of 1317 SSc patients and 3113 controls were included in the study. First we analyzed an initial case–control set of 332 SSc patients (242 with lcSSc and 90 with dcSSc) and 1296 healthy controls of Spanish Caucasian ancestry. Additionally, five independent replication cohorts were analyzed (Dutch: 101 lcSSc, 30 dcSSc and 893 controls; German: 153 lcSSc, 117 dcSSc and 227 controls; North American: 30 lcSSc, 53 dcSSc and 77 controls; Italian: 259 lcSSc, 92 dcSSc and 362 controls; Swedish: 111 lcSSc, 39 dcSSc and 285 controls).

All the patients fulfilled the 1980 American College of Rheumatology (ACR) classification criteria for SSc (28). In addition, patients were classified as having limited or diffuse SSc. When patients with SSc have cutaneous involvement distal from elbows, and knees they fulfilled definitions for limited scleroderma (29). Those SSc patients with cutaneous changes proximal from elbows and knees were classified as having diffuse SSc (30).

In addition, the following clinical data were collected for ascertainment of SSc clinical phenotype; age, gender, disease duration and presence of SSc specific auto-antibodies, anti-topoisomerase (Anti-Scl70) and anti-centromere (ACA). Lung involvement was assessed according to the international guidelines (31). Pulmonary fibrosis was assessed by a computed tomography scan. Restrictive syndrome and diffusion capacity of the lungs was defined as a FVC < 75% of the predicted value and a diffusion capacity for carbon monoxide (DLCO) of less than 75% of predicted (based on age, sex, height and ethnic origin). The main clinical features of the SSc patients from all the analyzed case sets are summarized in **table 4**.

The control population consisted in unrelated healthy individuals recruited in the same geographical region as SSc patients and matched by age, sex and ethnicity with the SSc patients groups.

The study was approved by local ethical committees from all the participating centers. Both patients and controls were included in the study after written informed consent.

Table 4. Main clinical features of SSc patients from the Spanish and the five replication cohorts

Phenotype	Spanish	Dutch	German	American	Swedish	Italian
Female (%)	87	72	87	100	77	95
Limited phenotype (%)	73	65	56.7	34	66	73
ACA positivity (%)	49	26	40	17	20	40
Anti-Scl70 positivity (%)	19	21	23	29	14.5	24
Pulmonary fibrosis on CT scan (%)	32	40	36	–	45	36.2
Low FVC (<75% predicted) (%)	28	24	18	–	35	18
Low DLCO (<75% predicted) (%)	45	20	50	–	24	50

STAT4 genotyping

DNA from patients and controls was obtained using standard methods. As *STAT4* genetic marker, we selected the rs7574865 polymorphisms, since this is the tagger SNP of the haplotype block of *STAT4* intron 3 associated with autoimmunity, and at the same time the genetic variant showing the strongest association with AIDs. Samples were genotyped for the rs7574865 polymorphism by Taqman 5′-allelic discrimination assay technology using a Pre-designed SNP Genotyping Assaysprovided by Applied Biosystems (Part number: C__29882391_10, Foster City, CA, USA). The PCR reaction was performed in a total volume of 5 µl with the following amplification protocol: denaturation at 92°C for 10 min, followed by 40 cycles of denaturation at 92°C for 15 s and annealing and extension at 60°C for 1:00 min. Post-PCR, the genotype of each sample was automatically attributed by measuring the allele-specific fluorescence in the ABI Prism 7900 Sequence Detection System, using the SDS 2.3 software for allele discrimination (Applied Biosystems, Foster City, CA, USA).

All samples were genotyped in the same center to avoid genotyping inconsistencies and to verify the genotyping consistency, alleatory samples were genotyped twice showing 99% identical genotypes.

Statistical analysis

We tested HWE for each case–control set by using the program FINETI. Significance was calculated by 2x2 contingency tables and Fisher's exact test, to obtain *p* values, OR and 95% CI by using Statcalc software (Epi Info 2002; Centers for Disease Control and Prevention, Atlanta, GA, USA). P-values below 0.05 were considered as statistically significant. The analysis of the combined data from all populations was performed using the Stats Direct software. First, homogeneity of OR among cohorts was calculated using Breslow-Day and Woolf Q methods. We then performed a calculation of the pooled OR under a fixed-effects model (Mantel–Haenszel meta-analysis) or random effects (DerSimonian-Laird) when necessary.

The estimation of the power of the study was performed using the Quanto v 0.5 software (Department of Preventive Medicine, University of Southern California, CA, USA). For the pooled analysis of lcSSc (n = 841) and considering a medium MAF of 0.25, our study reach a 93% power to detect the effect of a polymorphism at an OR of 1.3 similar to that observed for the rs7574865 T allele in previous studies (11,12). Under the same conditions, the estimation of the power for the pooled analysis of dcSSc that included a total of 421 patients was 67%.

References

1. Jimenez, S.A. and Derk, C.T. (2004) Following the molecular pathways toward an understanding of the pathogenesis of systemic sclerosis. Ann. Intern. Med., 140, 37 – 50.

2. Agarwal, S.K., Tan, F.K. and Arnett, F.C. (2008) Genetics and genomics studies in scleroderma (systemic sclerosis). Rheum. Dis. Clin. N. Am., 34, 17–40.

3. Merriman, T.R. and Pearce, S.H. (2006) Genetic progress towards the molecular basis of common autoimmunity. Discov. Med., 6, 40–45.

4. Gu S, Y., Kong, J., Cheema, G.S., Keen, C.L., Wick, G. And Gershwin, M.E. (2008) The immunobiology of systemic sclerosis. Semin. Arthritis Rheum., 38, 132–160.

5. Peng, S.L. (2008) Transcription factors in autoimmune diseases. Front. Biosci., 13, 4218–4240.

6. Lim, C.P. and Cao, X. (2006) Structure, function, and regulation of STAT proteins. Mol. Biosyst., 2, 536–550.

7. Ross, J.A., Nagy, Z.S., Cheng, H., Stepkowski, S.M. and Kirken, R.A. (2007) Regulation of T cell homeostasis by JAKs and STATs. Arch. Immunol. Ther. Exp. (Warsz.), 55, 231–245.

8. Watford, W.T., Hissong, B.D., Bream, J.H., Kanno, Y., Muul, L. And O'shea, J.J. (2004) Signaling by IL-12 and IL-23 and the immunoregulatory roles of STAT4. Immunol. Rev., 202, 139–156.

9. Mathur, A.N., Chang, H.C., Zisoulis, D.G., Stritesky, G.L., Yu, Q., O'Malley, J.T., Kapur, R., Levy, D.E., Kansas, G.S. and Kaplan, M.H. (2007) Stat3 and Stat4 direct development of IL-17-secreting Th cells. J. Immunol., 178, 4901–4907.

10. Deleuran, B. and Abraham, D.J. (2007) Possible implication of the effector CD4þ T-cell subpopulation TH17 in the pathogenesis of systemic scleroderma. Nat. Clin. Pract. Rheumatol., 3, 682–683.

11. Remmers, E.F., Plenge, R.M., Lee, A.T., Graham, R.R., Hom, G., Behrens, T.W., de Bakker, P.I., Le, J.M., Lee, H.S., Batliwalla, F. et al. (2007) STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N. Engl. J. Med., 357, 977 – 986.

12. Orozco, G., Alizadeh, B.Z., Delgado-Vega, A.M., Gonzalez-Gay, M.A., Balsa, A., Pascual-Salcedo, D., Fernandez-Gutierrez, B., Gonzalez-Escribano, M.F., Petersson, I.F., van Riel, P.L. et al. (2008) Association of STAT4 with rheumatoid arthritis: a replication study in three European populations. Arthritis. Rheum., 58,1974 – 1980.

13. Palomino-Morales, R.J., Rojas-Villarraga, A., Gonzalez, C.I., Ramirez, G., Anaya, J.M. and Martin, J. (2008) STAT4 but not TRAF1/C5 variants influence the risk of developing rheumatoid arthritis and systemic lupus erythematosus in Colombians. Genes. Immun., 9, 379–982.

14. Sigurdsson, S., Nordmark, G., Garnier, S., Grundberg, E., Kwan, T., Nilsson, O., Eloranta, M.L., Gunnarsson, I., Svenungsson, E., Sturfelt, G. et al. (2008) A risk haplotype of STAT4 for systemic lupus erythematosus is over-expressed, correlates with anti-dsDNA and shows additive effects with two risk alleles of IRF5. Hum. Mol. Genet., 17, 2868–2876.

15. Zervou, M.I., Mamoulakis, D., Panierakis, C., Boumpas, D.T. and Goulielmos, G.N. (2008) STAT4: a risk factor for type 1 diabetes? Hum. Immunol., 69, 647–650.

16. Martinez, A., Varade, J., Marquez, A., Cenit, M.C., Espino, L., Perdigones, N., Santiago, J.L., Fernandez-Arquero, M., de la Calle, H., Arroyo, R. et al. (2008) Association of the STAT4 gene with increased susceptibility for some immune-mediated diseases. Arthritis. Rheum., 58, 2598 – 2602.

17. Korman, B.D., Alba, M.I., Le, J.M., Alevizos, I., Smith, J.A., Nikolov, N.P., Kastner, D.L., Remmers, E.F. and Illei, G.G. (2008) Variant form of STAT4 is associated with primary Sjogren's syndrome. Genes. Immun., 9, 267 – 270.

18. Denton, C.P., Black, C.M. and Abraham, D.J. (2006) Mechanisms and consequences of fibrosis in systemic sclerosis. Nat. Clin. Pract. Rheumatol., 2, 134–144.

19. Kavvoura, F.K. and Ioannidis, J.P. (2008) Methods for meta-analysis in genetic association studies: a review of their potential and pitfalls. Hum. Genet., 123, 1–14.

20. Ioannidis, J.P. (2008) Effect of formal statistical significance on the credibility o f observational associations. Am. J. Epidemiol., 168, 374 – 383.

21. Bettelli, E., Oukka, M. and Kuchroo, V.K. (2007) T(H)-17 cells in the circle of immunity and autoimmunity. Nat. Immunol., 8, 345–350.

22. Sato, S., Hanakawa, H., Hasegawa, M., Nagaoka, T., Hamaguchi, Y., Nishijima, C., Komatsu, K., Hirata, A. and Takehara, K. (2000) Levels of interleukin 12, a cytokine of type 1 helper T cells, are elevated in sera from patients with systemic sclerosis. J. Rheumatol., 27, 2838 – 2842.

23. Komura, K., Fujimoto, M., Hasegawa, M., Ogawa, F., Hara, T., Muroi, E., Takehara, K. and Sato, S. (2008) Increased serum interleukin 23 in patients with systemic sclerosis. J. Rheumatol., 35, 120–125.

24. Gregersen, P.K., Lee, H.S., Batliwalla, F. and Begovich, A.B. (2006) PTPN22: setting thresholds for autoimmunity. Semin. Immunol., 18, 214–223.

25. Gregersen, P.K. and Behrens, T.W. (2006) Genetics of autoimmune diseases—disorders of immune homeostasis. Nat. Rev. Genet., 7, 917 – 928.

26. Cho, J.H. (2008) The genetics and immunopathogenesis of inflammatory bowel disease. Nat. Rev. Immunol., 8, 458 – 466.

27. Brionez, T.F. and Reveille, J.D. (2008) The contribution of genes outside the major histocompatibility complex to susceptibility to ankylosing spondylitis. Curr. Opin. Rheumatol., 20, 384 – 391.

28. Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. (1980) Preliminary criteria for the classification of systemic sclerosis (scleroderma). Arthritis. Rheum., 23, 581–590.

29. LeRoy, E.C. and Medsger, T.A. Jr (2001) Criteria for the classification of early systemic sclerosis. J. Rheumatol., 28, 1573–1576.

30. LeRoy, E.C., Black, C., Fleischmajer, R., Jablonska, S., Krieg, T., Medsger, T.A. Jr, Rowell, N. and Wollheim, F. (1988) Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. J. Rheumatol., 15, 202–205.

31. Matucci-Cerinic, M., D'Angelo, S., Denton, C.P., Vlachoyiannopoulos, P. and Silver, R. (2003) Assessment of lung involvement. Clin. Exp. Rheumatol., 21, S19–S23.

Chapter 3

The *FAS-670A>G* polymorphism influences susceptibility to systemic sclerosis phenotypes

Broen J, Gourh P, Rueda B, Coenen M, Mayes M, Martin J, Arnett FC, Radstake TR; European Consortium on Systemic Sclerosis Genetics.

Arthritis and Rheumatism 2009

Abstract

Objective: To investigate the possible role of the *FAS* −670A>G functional polymorphism in the genetic predisposition to systemic sclerosis (SSc) susceptibility or clinical phenotype.

Methods: A total of 2,900 SSc patients and 3,186 healthy controls were included in this study. We analyzed the genotype and allele frequencies of the *FAS* −670A>G polymorphism in 9 distinct ethnic cohorts, including 6 cohorts of European ancestry (a Spanish cohort of 228 SSc patients and 265 controls, a Dutch cohort of 203 SSc patients and 277 controls, a German cohort of 313 SSc patients and 247 controls, an Italian cohort of 323 SSc cases and 89 controls, a British cohort of 269 SSc patients, and a Swedish cohort of 182 patients) and 3 distinct ethnic cohorts from the US (a cohort of 1,047 white patients and 692 controls, a cohort of 159 Hispanic patients and 137 controls, and a cohort of 176 black SSc patients and 194 controls). Genotyping was performed using a TaqMan 5′ allelic discrimination assay.

Results: In the British, Italian, and American white cohorts we observed an association of the *FAS* −670G allele with limited cutaneous SSc (lcSSc) (odds ratios [ORs] 1.25, 1.43, and 1.18, respectively). A meta-analysis comprising all 9 cohorts revealed an association of both the *FAS* −670G allele (OR 1.10) and the FAS −670GG genotype (OR 1.13) with the lcSSc phenotype. In a meta-analysis including only white subjects, both the *FAS* −670G allele and the *FAS* −670GG genotype remained associated with lcSSc (allele OR 1.12; genotype OR 1.16). In addition, a recessive model of the −670GG genotype exhibited a strong association with SSc, lcSSc, and anticentromere antibody–positive lcSSc (OR 1.23, OR 1.33, and OR 1.45, respectively).

Conclusion: Our data show that the *FAS* −670A>G polymorphism plays a role in lcSSc susceptibility. A similar trend has been observed in other autoimmune diseases.

Introduction

Systemic sclerosis (SSc; scleroderma) is a connective tissue disease in which patients develop extensive fibrosis of the skin and internal organs. Based on the extent of skin involvement, the disease can be classified as limited cutaneous SSc (lcSSc) or diffuse cutaneous SSc (dcSSc) (1). In the early stage of SSc, perivascular infiltrations of immune cells are observed, among which T cells and antigen-presenting cells are key players (2).

Intriguingly, some T cell subsets in patients with SSc exhibit a decreased response to activation-induced cell death and apoptosis compared with healthy controls (3). One of the main activators of apoptosis in T cells is soluble Fas, which has been found to be elevated in SSc serum (4). The *FAS* gene has been described as an "autogene," because its dysregulated function contributes to various autoimmune diseases. A common single-nucleotide polymorphism (SNP), *FAS* −670A>G (rs1800682), occurring at the binding sequence of the interferon-γ activation site, has been reported to confer susceptibility to systemic lupus erythematosus, multiple sclerosis, sarcoidosis, and autoimmune hepatitis (5–8). Recently, the *FAS* −670A allele was found to be significantly more frequent in a cohort of 350 Italian SSc patients compared with healthy controls; additionally, the *FAS* −670AA genotype influenced the predisposition to SSc in general and to both lcSSc and dcSSc (9). Insight into the potential role of Fas in SSc pathogenesis would greatly facilitate our understanding of the disease. Therefore, we studied the *FAS* −670A>G polymorphism in 9 large independent SSc case–control series comprising 2,900 SSc patients and 3,186 controls of multiple races.

Methods

Patients and controls.

DNA samples from European subjects were provided by the European Consortium on Systemic Sclerosis Genetics (acknowledgements). The study population was composed of 2,900 SSc patients and 3,186 healthy controls matched by geographic region, age, and sex. Six case–control sets were of European ancestry (a Spanish cohort of 228 SSc patients and 265 controls, a Dutch cohort of 203 SSc patients and 277 controls, a German cohort of 313 SSc patients and 247 controls, an Italian cohort of 323 SSc cases and 89 controls, a British cohort of 269 SSc patients, and a Swedish cohort of 182 patients). The genotype frequency in the 351 Swedish and 934 British controls was derived from literature reports (10, 11). Additionally, 3 distinct ethnic cohorts resident in the US were considered in the −670A>G genotyping (1,047 American white SSc patients and 692 matched controls, 159 American Hispanic SSc patients and 137 matched controls, and 176 American black SSc patients and 194 controls). All patients fulfilled the American College of Rheumatology (formerly, the American Rheumatism Association) 1980 classification criteria for SSc (12). The local ethics committee from each center approved the study. Patients and controls provided written informed consent before enrollment in the study. All patients included in this study were classified as having lcSSc or dcSSc, using the criteria proposed by LeRoy et al (1). In addition, the presence or absence of antibodies (anti–topoisomerase I (ATA) and anti-centromere (ACA)) was recorded (**table 1**).

Genotyping of the FAS −670A>G polymorphism.

DNA samples from patients and controls were genotyped for the *FAS* −670A>G polymorphism (rs1800682) with a TaqMan SNP genotyping assay using the ABI 7500/7900HT real time thermocycler according to the protocol recommended by the manufacturer (Applied Biosystems, Foster City, CA). Automated allele calling was performed using SDS 2.3 software from Applied Biosystems. Multiple positive Centre d'Etude du Polymorphisme Humain DNA samples from Coriell Institute for Medical Research (Camden, NJ) and negative controls were used in each genotyping assay.

Table 1. Demographic and clinical characteristics of the 9 SSc cohorts included in the present study

Phenotype	Population								
	The Netherlands	Spain	Germany	Sweden	Italy	UK	American white	American Hispanic	American black
n	203	228	313	182	323	269	1,047	159	176
Female, %	71.7	83.8	87.4	76.9	95.5	81.3	73.9	74.4	78.2
Age, mean ± SD years	56 ± 13	58 ± 13	56 ± 12	56 ± 15	55 ± 13	59 ± 12	NA	NA	NA
Disease duration, mean ± SD months	133 ± 87	144 ± 90	110 ± 109	81 ± 73	141 ± 138	154 ± 91	NA	NA	NA
lcSSc, %	77.5	70.0	52.5	82.1	70.7	74.6	60.8	41.1	32.5
ATA positive, %	20.3	18.4	24.7	16.7	46.9	6.3	16.5	22.0	23.0
ACA positive, %	26.8	46.7	40.4	26.1	48.1	38.5	29.0	7.0	16.0

Statistical analysis.

Hardy-Weinberg equilibrium was tested with the program FINETI. Significance levels were calculated with 2 × 2 contingency tables and Fisher's exact test, using SPSS 16.0. P values less than 0.05 were considered significant. Homogeneity of odds ratios (ORs) among cohorts was calculated using the Breslow-Day and Woolf Q methods, and pooled ORs were calculated under a fixed-effects model (Mantel-Haenszel meta-analysis). Ninety-five percent confidence intervals (95% CIs) were calculated. The statistical power of the study was estimated using Quanto software, version 0.5. For the power calculation in the pooled analysis of white lcSSc patients (n = 1,552), we considered a minor allele frequency of 0.46, as reported in the HapMap data. The study reaches a power of 80% to detect the effect of a polymorphism at an OR of 1.26 in a recessive model and 1.25 in a dominant

model. Under the same conditions, the estimation of the power for the pooled analysis of dcSSc (n = 811) is 80% to detect an OR of 1.37 in a recessive model and 1.36 in a dominant model. Although the clinical features tested are not truly independent, we applied the Bonferroni correction for multiple testing. Corrected *p* values less than 0.02 were considered significant.

Results

The FAS −670G allele is associated with lcSSc in 3 cohorts.

The *FAS* −670G allele frequencies in the controls enrolled in the study cohorts were consistent with frequencies reported in the HapMap project. To evaluate the effect of the *FAS* −670A>G polymorphism on SSc susceptibility, we compared genotype and allele frequencies in the healthy controls with those in the total SSc group and the lcSSc and dcSSc subgroups. Presence of the *FAS* −670G allele was found to be associated with lcSSc in 3 cohorts: the British cohort (OR 1.25 [95% CI 1.00–1.60], $p = 0.049$), the Italian cohort (OR 1.43 [95% CI 1.00–2.07], $p = 0.045$), and the American white cohort (OR 1.18, [95% CI 1.01-1.39], $p = 0.036$). In addition, in the American white cohort, an association of the *FAS* −670GG genotype was observed for lcSSc (OR 1.26 [95% CI 1.02–1.55], $p = 0.017$) and SSc overall (OR 1.18 [95% CI 0.98–1.41], $p = 0.022$) (**table 2**).

A meta-analysis with the 9 distinct ethnic populations confirms that the *FAS* −670G allele and FAS −670GG genotype are more frequent in patients with lcSSc.

To validate the associations observed in the 3 independent populations (**table 2**), we performed a meta-analysis under a fixed-effects model (Mantel-Haenszel) for both the total SSc cohort and the SSc subtypes. No statistically significant deviation in heterogeneity between cohorts was observed ($p = 0.575$, $p = 0.280$, and $p = 0.385$, controls versus all SSc, lcSSc, and dcSSc, respectively). The *FAS* −670G allele and the *FAS* −670GG genotype were found to be associated with lcSSc (Table 3) (allele OR 1.10 [95% CI 1.01–1.21], $p = 0.036$; genotype OR 1.13 [95% CI 1.01–1.27], $p = 0.036$).

Table 2. Genotype and allele distributions of the FAS −670A>G polymorphism

Phenotype	n	AA	AG	GG	*P* vs. controls	Allele, *P* vs. controls	−670GG recessive, *P* vs. controls
Italy							
SSc	323	0.27	0.46	0.27	0.196	0.078	0.084
lcSSc	213	0.26	0.44	0.30	0.105	0.045	0.156
ACA+ lcSSc	102	0.27	0.45	0.28	0.223	0.106	0.089
dcSSc	84	0.26	0.52	0.21	0.627	0.358	0.568
Control	89	0.33	0.49	0.18	–	–	–
Sweden							
SSc	182	0.32	0.47	0.20	0.413	0.366	0.899
lcSSc	117	0.33	0.46	0.21	0.463	0.52	0.896
ACA+ lcSSc	41	0.32	0.42	0.27	0.421	0.905	0.373
dcSSc	50	0.34	0.48	0.18	0.587	0.361	0.646
Control‡	351	0.27	0.53	0.21	–	–	–
Spain							
SSc	228	0.31	0.42	0.27	0.046	0.652	0.290
lcSSc	154	0.33	0.41	0.27	0.046	0.545	0.358
ACA+ lcSSc	78	0.32	0.42	0.26	0.213	0.59	0.582
dcSSc	58	0.22	0.48	0.29	0.557	0.412	0.280
Control	265	0.24	0.53	0.23	–	–	–
Germany							
SSc	313	0.32	0.45	0.24	0.106	0.446	0.072
lcSSc	163	0.33	0.42	0.25	0.073	0.465	0.057
ACA+ lcSSc	118	0.32	0.42	0.26	0.073	0.400	0.049
dcSSc	121	0.31	0.48	0.21	0.647	0.817	0.450
Control	247	0.30	0.53	0.17	–	–	–
UK							
SSc	269	0.28	0.49	0.23	0.731	0.465	0.702
lcSSc	172	0.22	0.52	0.26	0.104	0.049	0.295
ACA+ lcSSc	78	0.27	0.47	0.26	0.714	0.417	0.451
dcSSc	59	0.31	0.56	0.14	0.282	0.346	0.128
Control‡	934	0.30	0.48	0.22	–	–	–
The Netherlands							
SSc	203	0.28	0.50	0.23	0.89	0.713	0.631
lcSSc	130	0.28	0.49	0.24	0.985	0.861	0.878
ACA+ lcSSc		0.15	0.50	0.35	0.175	0.060	0.158
dcSSc	50	0.36	0.52	0.12	0.120	0.048	0.051
Control	277	0.27	0.48	0.25	–	–	–

Phenotype	n	AA	AG	GG	*P* vs. controls	Allele, *P* vs. controls	−670GG recessive, *P* vs. controls
American white							
SSc	1,047	0.29	0.49	0.23	0.022	0.086	0.007
lcSSc	603	0.27	0.49	0.24	0.017	0.032	0.004
ACA+ lcSSc	341	0.28	0.49	0.24	0.07	0.124	0.023
dcSSc	389	0.32	0.47	0.21	0.091	0.819	0.120
Control	692	0.29	0.54	0.18	–	–	–
American Hispanic							
SSc	159	0.21	0.50	0.29	0.863	0.669	0.851
lcSSc	62	0.21	0.48	0.31	0.873	0.852	0.919
ACA+ lcSSc	46	0.24	0.48	0.28	0.704	0.541	0.830
dcSSc	89	0.20	0.51	0.29	0.934	0.779	0.909
Control	137	0.18	0.52	0.30	–	–	–
American black							
SSc	176	0.09	0.39	0.53	0.718	0.699	0.519
lcSSc	53	0.11	0.43	0.45	0.675	0.438	0.588
ACA+ lcSSc	44	0.14	0.41	0.46	0.457	0.359	0.629
dcSSc	110	0.07	0.36	0.57	0.412	0.275	0.191
Control	194	0.08	0.43	0.50	–	–	–

*Patients with an undifferentiated skin phenotype were excluded from phenotype analysis.

† Control frequencies derived from ref.10. ‡ Control frequencies derived from reference11.

The effect of the FAS −670G allele on lcSSc is most prominent in white ACA-positive lcSSc patients.

Although we observed no significant deviation in heterogeneity between the 9 cohorts, it was clear that there was a large difference in the frequency of the *FAS* −670G allele in the total group of white subjects compared with the black American and Hispanic American subjects. In contrast to the white population, in both the black and Hispanic Americans the *FAS* −670G allele was the major allele. For this reason we performed an additional meta-analysis solely on the 7 white study populations. We found that both the *FAS* −670G allele and the *FAS* −670GG genotype were associated with lcSSc (allele OR 1.12 [95% CI 1.02–1.24], $p = 0.02$; genotype OR 1.16 [95% CI 1.03–1.31], $p = 0.017$). In addition, we analyzed the subgroup of lcSSc patients who were positive for ACA. In all white populations we observed that the frequency of the *FAS* −670GG genotype was higher in this subgroup than in controls. Although the difference did not reach significance in any

single population, the meta-analysis revealed a significant effect of both the minor allele (OR 1.16 [95% CI 1.03–1.31], $p = 0.01$) and the −670GG genotype (OR 1.21 [95% CI 1.03–1.41], $p = 0.017$) (**table 3**).

Implementation of the FAS −670GG genotype in a recessive model reveals a strong association with SSc and lcSSc in whites.

When evaluating the frequencies of the *FAS* genotypes in cases and controls, we observed a higher frequency of the −670GG in the lcSSc patients, whereas the −670AG genotype was more frequent in controls. This finding, taken together with reports of a strong effect of the *FAS* −670GG genotype on *FAS* expression levels (9, 13), led us to hypothesize that a recessive effect of the *FAS* −670GG genotype in SSc susceptibility should be considered. To investigate this, we compared the proportion of white study subjects with the *FAS* −670GG genotype versus the proportion with the FAS −670AG or the *FAS* −670AA genotype (grouped together). In the meta-analysis we observed a strong association of the *FAS* −670GG genotype with SSc overall (OR 1.23 [95% CI 1.07–1.41], $p = 0.004$), lcSSc (OR 1.33 [95% CI 1.14–1.56], $p = 0.0003$), and ACA-positive lcSSc (OR 1.45 [95% CI 1.19–1.76], $p = 0.0002$) (**table 3**).

Table 3. Overall genotype frequencies and Mantel-Haenszel meta-analysis under a fixed-effects model

		Genotype				−670GG recessive,	Allele		
Phenotype	n	AA	AG	GG	*P* vs. controls	*P* vs. controls	A	G	*P* vs. controls
All									
SSc	2,900	0.27	0.47	0.26	0.173	0.005	0.51	0.49	0.172
lcSSc	1,667	0.27	0.47	0.26	0.036	0.001	0.51	0.49	0.036
ACA+ lcSSc	770	0.28	0.46	0.26	0.032	0.0017	0.50	0.50	0.056
dcSSc	1,010	0.27	0.47	0.25	0.925	0.322	0.51	0.49	0.862
Control	3,186	0.27	0.50	0.23	–	–	0.52	0.48	–
Caucasians									
SSc	2,565	0.29	0.47	0.24	0.147	0.004	0.53	0.47	0.145
lcSSc	1,552	0.28	0.47	0.25	0.02	0.0003	0.51	0.49	0.017
ACA+ lcSSc	686	0.29	0.46	0.25	0.017	0.0002	0.51	0.49	0.01
dcSSc	811	0.31	0.49	0.20	0.714	0.548	0.55	0.45	0.718
Control	2,855	0.29	0.51	0.21	–	–	0.54	0.46	–

Discussion

The aim of this study was to validate the results of a previous study depicting *FAS* −670A>G as a novel underlying genetic variant influencing SSc susceptibility (9). In the present investigation, we showed that the *FAS* −670A>G variant plays a significant role in susceptibility to SSc, lcSSc, and ACA-positive lcSSc in white populations. After Bonferroni correction for multiple comparisons, most associations remained statistically significant, underscoring their soundness.

The −670 variant in the *FAS* promoter influences gene expression, with the G variant disrupting the interferon-γ binding site for the transcription factor *STAT1*. It has been shown that healthy subjects who are homozygous for the −670A major allele have higher levels of *FAS* expression than those who are homozygous for the −670G variant (13). It is tempting to speculate that the up-regulation of soluble FAS observed in SSc patients is an endeavor of the immune system to diminish the autoreactive immune processes observed in SSc. From this standpoint, FAS levels would be less able to increase through the STAT1 signaling cascade in individuals carrying the −670GG genotype, and these individuals would be exposed to autoreactive T cell clones at an increased rate. However, we did not observe a significant effect of the *FAS* −670G allele on dcSSc. This is consistent with a previous report describing a strong effect of a *STAT4* polymorphism on lcSSc susceptibility but not on dcSSc (14). It suggests that there are marked differences in the genetic background underlying these disease subtypes.

It should be noted that our findings contradict results previously reported by Liakouli et al, from a study of Italian subjects (9). In that study, the *FAS* −670A allele was found to be associated with SSc susceptibility. We suggest that an underlying factor contributing to this apparent discrepancy might be that the frequency of the minor G allele in Liakouli and colleagues' Italian control group (0.52) differs from the allele frequency reported in the International HapMap project, as well as that found in other Italian studies investigating the *FAS* −670A>G polymorphism (15, 16) and that observed in the present study. The frequency observed by Liakouli et al may thus not be representative of the general Italian population. On the other hand, conflicting results in rheumatoid arthritis and

primary Sjögren's syndrome have been described as well. Considering the power of the study, it is unlikely that our findings were biased by a type I error.

Associations of the *FAS* −670A>G polymorphism with other autoimmune diseases have already been documented. Therefore, it is conceivable that this polymorphism, together with others, is part of a genetic framework that renders an individual susceptible to a breach of immune tolerance. On this framework, additional genetic and environmental factors determine the characteristics of the full-blown autoimmune disease. Fortunately, we are in an era in which genetic research is evolving at an astonishing pace, and with the use of whole-genome association studies and subsequent meta-analyses, more of the genetic factors underlying SSc etiology will certainly be illuminated.

References

1. LeRoy EC, Black C, Fleischmajer R, Jablonska S, Krieg T, Medsger TA Jr, et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. J Rheumatol 1988;15:202–5.

2. Roumm AD, Whiteside TL, Medsger TA Jr, Rodnan GP. Lym- phocytes in the skin of patients with progressive systemic sclerosis: quantification, subtyping, and clinical correlations. Arthritis Rheum 1984;27:645–53.

3. Cipriani P, Fulminis A, Pingiotti E, Marrelli A, Liakouli V, Perricone R, et al. Resistance to apoptosis in circulating '/' and '/' T lymphocytes from patients with systemic sclerosis. J Rheu- matol 2006;33:2003–14.

4. Wetzig T, Petri JB, Mittag M, Haustein UF. Serum levels of soluble Fas/APO-1 receptor are increased in systemic sclerosis. Arch Dermatol Res 1998;2904:187–90.

5. Kanemitsu S, Ihara K, Saifddin A, Otsuka T, Takeuchi T, Na- gayama J, et al. A functional polymorphism in fas (CD95/APO-1) gene promoter associated with systemic lupus erythematosus. J Rheumatol 2002;29:1183–8.

6. Wasfi YS, Silveira LJ, Jonth A, Hokanson JE, Fingerlin T, Sato, et al. Fas promoter polymorphisms: genetic predisposition to sarcoidosis in African-Americans. Tissue Antigens 2008;721: 39–48.

7. Agarwal K, Czaja AJ, Donaldson PT. A functional Fas promoter polymorphism is associated with a severe phenotype in type 1 autoimmune hepatitis characterized by early development of cirrhosis. Tissue Antigens 2007;69:227–35.

8. Van Veen T, Kalkers NF, Crusius JB, van Winsen L, Barkhof F, Jongen PJ, et al.The FAS-670 polymorphism influences susceptibility to multiple sclerosis. J Neuroimmunol 2002;128:95–100.

9. Liakouli V, Manetti M, Pacini A, Tolusso B, Fatini C, Toscano A, et al. The -670G_A polymorphism in the FAS gene promoter region influences the susceptibility to systemic sclerosis. Ann Rheum Dis 2009;68:584–90.

10. Zhang H, Sun XF, Synnerstad I, Rosdahl I. Importance of FAS-1377, FAS-670, and FASL-844 polymorphisms in tumor onset, progression, and pigment phenotypes of Swedish patients with melanoma: a case-control analysis. Cancer J 2007;13:233–7.

11. Sibley K, Rollinson S, Allan JM, Smith AG, Law GR, Roddam PL, et al. Functional FAS promoter polymorphisms are associated with increased risk of acute myeloid leukemia. Cancer Res 2003; 6315:4327–30.

12. Subcommittee for Scleroderma Criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Commit- tee. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Arthritis Rheum 1980;23:581–90.

13. Mahfoudh W, Bel H Jr., Romdhane A, Chouchane L. A polymorphism in FAS gene promoter correlated with circulating soluble FAS levels. Int J Immunogenet 2007;34:209–12.

14. Rueda B, Broen J, Simeon C, Hesselstrand R, Diaz B, Suarez H, et al. The STAT4 gene influences the genetic predisposition to systemic sclerosis phenotype. Hum Mol Genet 2009;18:2071–7.

15. Nasi M, Pinti M, Bugarini R, Troiano L, Lugli E, Bellodi C, et al. Genetic polymorphisms of Fas (CD95) and Fas ligand (CD178) influence the rise in CD4_ T cell count after antiretroviral therapy in drug-naive HIV-positive patients. Immunogenetics 2005;57: 628–35.

16. Andreoli V, Nicoletti G, Romeo N, Condino F, La Russa A, Liguori M, et al. Fas antigen and sporadic Alzheimer's disease in Southern Italy: evaluation of two polymorphisms in the TNFRSF6 gene. Neurochem Res 2007;32:1445–9.

Chapter 4

The functional polymorphism *844 A>G* in the *Fc-alpha-RI* (CD89) gene does not contribute to systemic sclerosis or rheumatoid arthritis susceptibility.

Broen JC, Coenen MJ, Rueda B, Witte T, Padyukov L, Klareskog L, Hesselstrand R, Wuttge DM, Simeon C, Ortego-Centeno N, González-Gay MA, Pros A, Hunzelman N, Riemekasten G, Kreuter A, Vonk M, Scorza R, Beretta L, Airò P, van Riel PL, Kimberly R, Martin J, Edberg J, Radstake TR.

The Journal of Rheumatology 2011

Abstract

Objective: To investigate the role of the *FCAR1* 844A>G functional polymorphism in the genetic predisposition to rheumatoid arthritis (RA) and systemic sclerosis (SSc) susceptibility.

Methods: The study population was composed of 1401 SSc patients, 642 RA patients and 1317 healthy controls. The *FCAR1 (CD89)* SNP rs16986050 was genotyped by Pyrosequencing.

Results: We observed no significant deviation of the genotype and allele frequencies in RA and SSc compared to controls. A meta-analysis, a recessive and dominant model yielded similar negative results.

Conclusion: Our data show that the *FCAR1* 844A>G polymorphism is not associated with SSc or RA susceptibility.

Introduction

Fc-receptors (FcR) play a pivotal role in linking humoral and cellular components of immunity by effectuating the recognition of antigens bound to immunoglobulins (Ig). Interestingly, there is a large body of evidence describing genetic variations in *FcR* that were found associated with a wide range of autoimmune pathologies, such as systemic lupus erythematosus (SLE) and rheumatoid arthritis (RA) (1). Although disease specific IgA is present in auto-immune disease, its impact remains to be the least scrutinized. IgA exhibits its immunoregulatory functions mainly by triggering cellular effector functions through the FC alpha receptor (FCAR) on the cell surface. FCAR1 (CD89) is the functionally most important IgA receptor and is expressed on various cells of the immune system (2). Serum IgA regulates secretion of IL1β, IL10, TNF-α, IL6 and IL1RA. To mediate these effects, the FcR-associated signal-transducing transmembrane gamma-chain (FCGR) needs to be present in complex with FCAR1 (3). Intriguingly, a polymorphism (844A>G, rs16986050) in the coding region of *FCAR1* leads to an amino-acid change of Serine 248 to Glycine (S248G) which potentates IL6 production and is able to induce cytokine release in the absence of the *FCGR* chain. (4). This variant was found to be enriched in two SLE populations compared to healthy controls (5). In other auto-immune diseases specific IgA antibodies are also present. For instance, anti-topoisomerase IgA and rheumatoid factor (RF) IgA and anti-CCP IgA are present in systemic sclerosis (SSc) and RA (6;7). Although the exact role of IgA in these diseases remains to be elucidated, we hypothesized that an increased inflammatory response upon disease specific IgA binding caused by this polymorphism could contribute to these diseases. For this reason we investigated the frequency of this variant in SSc and RA patients.

Methods

Study population

The study population was composed of 1401 SSc patients, 642 Dutch RA patients and 1317 healthy controls derived from healthy blood donors, demographically, age and gender matched. Since SSc is a rare disease we composed a cohort of five case–control sets of European (Table 1). All patients fulfilled the American College of Rheumatology (ACR) 1980 classification criteria for SSc (8). The local ethics committee from each center approved the study. Patients and controls provided written informed consent. All patients included in this study were classified as having limited cutaneous SSc (lcSSc) or diffuse cutaneous SSc (dcSSc), using the criteria proposed by LeRoy et al. (9) Further information on clinical phenotype was recorded as well. Autoantibody testing was performed in every center separately using either ELISA or immunofluorescence microscopy (**table 1**). The study included RA patients who met the ACR 1987 revised criteria for RA (10) (**table 2**).

Table 1. Clinical characteristics of the SSc patients and controls included in this study

Population	Netherlands	Spain	Germany	Sweden	Italy
Number	143	231	422	161	444
Age (years, SD)	56 (13)	55 (13)	57(12)	53 (15)	54(13)
Disease duration (months,SD)	132 (87)	143 (90)	112 (109)	83 (73)	146 (138)
Female %	73.2	83.8	83.2	78.8	89.0
Limited phenotype%	74.8	73.7	61.2	72.0	74.4
Positivity ANA%	73.2	85.4	83.2	75.1	81.6
Positivity anti-topo%	22	N.A.	24.8	16.7	25.2
Positivity ACA%	26.8	46.7	41.3	26.1	47.1
Healthy Control Age (years, SD)	45 (12)	56 (14)	48 (9)	52 (11)	52 (14)
Healthy Control Female %	69.3	85.1	82.4	80.2	85.0

Genotyping

The *FCAR1* 844A>G variant was genotyped by pyrosequencing, congruent with previously published reports from our group, pyrosequencing reactions were performed according to the manufacturer's instructions on a PSQ-HS96A system (Biotage) (5, 11).

Table 2. Clinical characteristics of the rheumatoid arthritis patients included in this study

Number	642
Age (years, SD)	65 (13)
Female %	66.5
Positivity RF%	78
Positivity Anti-CCP%*	65

*status available in 155 patients

Statistical analysis

Significance levels were calculated with Fisher's exact test, P values below 0.05 were considered significant after Bonferroni adjustment. Homogeneity of odds ratios (ORs) was assessed with Breslow-Day statistics, pooled ORs were calculated under a fixed-effects model (Mantel-Haenszel). In the SSc population (n=1401), this study reaches a power of 80% to detect an OR of 1.21. The estimation of the power for the RA population (n=645) is 80% to detect an OR of 1.31.

Results

After genotyping no divergence in HWE was observed. We observed no significant deviation in genotype and allele frequencies of the *FCAR1* 844A>G polymorphism in Dutch RA patients compared to Dutch controls and when comparing ACPA+/- and RF+/- RA patients. No significant heterogeneity was detected between the five European populations, justifying a meta-analysis. Initially, we observed a divergence in genotype distribution in the Swedish (p =0.034) and Spanish (p =0.019) SSc populations. The same was observed for lcSSc susceptibility in both the Swedish (p =0.016) and Spanish (p =0.022) SSc population. The FCAR1 844G allele was associated with an increased risk to SSc, lcSSc and ACA+ SSc in the Swedish cohort (respectively; p =0.019, 0.030 and 0.032). The 844G allele was found less frequent in the Italian ACA+ SSc patients (p =0.042). However, after correction for multiple testing, no result remained significant. In addition, a meta-analysis, recessive and dominant models yielded similar negative results (**table 3** and data not shown).

Table 3. Genotype frequencies of the *FCAR1* 844A>G polymorphism in five European systemic sclerosis populations and a Dutch rheumatoid arthritis population.

Origin	Subtype	n total	AA%	AG%	GG%	*P* vs HC	A%	G%	*P* vs HC
Italy	**SSc**	444	0.76	0.22	0.02	0.484	0.87	0.13	0.277
	lcSSc	327	0.75	0.24	0.01	0.253	0.87	0.13	0.211
	dcSSc	112	0.74	0.21	0.05	0.409	0.85	0.15	0.903
	ACA+	144	0.80	0.19	0.01	0.123	0.90	0.10	**0.042**
	Anti-Topo+	146	0.71	0.27	0.01	0.631	0.85	0.15	0.923
	Control	362	0.73	0.25	0.02		0.85	0.15	
Sweden	**SSc**	161	0.58	0.30	0.12	**0.034**	0.73	0.27	**0.019**
	lcSSc	116	0.60	0.26	0.14	**0.016**	0.73	0.27	**0.030**
	dcSSc	45	0.53	0.40	0.07	0.252	0.73	0.27	0.107
	ACA+	44	0.55	0.32	0.13	0.055	0.70	0.30	**0.039**
	Anti-Topo+	25	0.64	0.28	0.08	0.712	0.78	0.22	0.591
	Control	165	0.67	0.29	0.04		0.81	0.19	
Spain	**SSc**	231	0.66	0.26	0.08	**0.019**	0.79	0.21	0.639
	lcSSc	157	0.66	0.25	0.09	**0.022**	0.79	0.21	0.794

Origin	Subtype	n total	AA%	AG%	GG%	*P* vs HC	A%	G%	*P* vs HC
	dcSSc	56	0.62	0.29	0.09	0.274	0.76	0.24	0.900
	ACA+	89	0.66	0.28	0.06	0.388	0.80	0.20	0.525
	Anti-Topo+	N.A.	N.A.	N.A.	N.A.		N.A.	N.A.	
	Control	250	0.60	0.36	0.04		0.78	0.22	
Germany	**SSc**	422	0.64	0.31	0.05	0.564	0.79	0.21	0.680
	lcSSc	261	0.68	0.29	0.03	0.532	0.82	0.18	0.612
	dcSSc	160	0.63	0.32	0.05	0.642	0.79	0.21	0.722
	ACA+	78	0.69	0.25	0.06	0.936	0.81	0.19	0.825
	Anti-Topo+	50	0.58	0.34	0.08	0.616	0.75	0.25	0.507
	Control	266	0.66	0.28	0.06		0.80	0.20	
Netherlands	**SSc**	143	0.69	0.28	0.03	0.434	0.83	0.17	0.769
	lcSSc	98	0.72	0.24	0.04	0.983	0.84	0.16	0.999
	dcSSc	33	0.61	0.39	0.00	0.09	0.80	0.20	0.484
	ACA+	34	0.76	0.24	0.00	0.455	0.88	0.12	0.383
	Anti-Topo+	33	0.70	0.27	0.03	0.860	0.83	0.17	0.862
	RA	642	0.71	0.26	0.03	0.593	0.84	0.16	0.582
	Control	274	0.72	0.24	0.04		0.84	0.16	
Total	**SSc**	1401	0.68	0.27	0.05	0.662	0.82	0.18	0.675
Mantel-Haenszel	**lcSSc**	959	0.68	0.26	0.06	0.842	0.81	0.19	0.825
	dcSSc	406	0.63	0.32	0.05	0.102	0.79	0.21	0.192
	ACA+	389	0.72	0.24	0.04	0.278	0.84	0.16	0.287
	Anti-Topo+	254	0.68	0.28	0.04	0.394	0.82	0.18	0.547
	Control	1317	0.68	0.28	0.04		0.82	0.18	

Discussion

In this study we show that a common polymorphism in the coding region of *FCAR1* is not associated with RA or SSc susceptibility in two large cohorts. Considering the power of our study to detect significant deviations in allele frequencies of the *FCAR1* 844G variant between cases and controls, it is unlikely that the lack of association is due to a type 2 error. This indicates that the *FCAR1* 844A>G polymorphism does not play a role in the susceptibility to RA and SSc and neither influences clinical phenotype. This is in contrast to the previous association of the *FCAR1* 844A>G polymorphism with SLE susceptibility (5). A number of polymorphisms have been found to influence susceptibility to RA as well as SSc and SLE (12, 13). These polymorphisms form therefore merely a genetic foundation for autoimmunity in general. Intriguingly, the polymorphism investigated in this study seems to be specific for SLE. An explanation for this can perhaps be found in the properties of IgA in these three conditions. In contrast to RA and SSc, mean total IgA has been found significantly elevated in SLE compared to controls, and IgA mediated inflammation may thus play a proportional larger role in SLE (14;15). IgA anti-CCP antibodies are present in 29% of the overall RA patients and in 47% of the RA patients with IgG anti-CCP antibodies. In SSc, IgA anti-topoisomerase antibodies have been described in 26.6% in a study containing 45 patients (15,16). However, this implies that in a subgroup of SSc and RA patients with high IgA antibody titers, which has been described in smoking RA patients, an effect of this variant may still be present (15). This data was not available from the cohorts described in this study In addition, autoantibody measurements have been performed in separate centers with either ELISA or immunofluorescence microscopy, recently has been described that caution is warranted when comparing outcomes from both techniques with each other (17). Future investigations into the role of IgA and the *FCAR1* gene might therefore benefit from taking this observation into account and focussing on subgroups with high IgA titers. Altogether, we could not demonstrate a role for the *FCAR1* 844A>G variant in SSc and RA.

References

1. Bournazos S, Woof JM, Hart SP, Dransfield I. Functional and clinical consequences of Fc receptor polymorphic and copy number variants. Clin Exp Immunol 2009; 157(2): 244-54.

2. Wines BD, Hogarth PM. IgA receptors in health and disease. Tissue Antigens 2006; 68(2):103-14.

3. Otten MA, van EM. The Fc receptor for IgA (FcalphaRI, CD89). Immunol Lett 2004; 92(1-2):23-31.

4. Jasek M, Manczak M, Sawaryn A, Obojski A, Wisniewski A, Luszczek W et al. A novel polymorphism in the cytoplasmic region of the human immunoglobulin A Fc receptor gene. Eur J Immunogenet 2004; 31(2): 59-62.

5. Wu J, Ji C, Xie F, Langefeld CD, Qian K, Gibson AW et al. FcalphaRI (CD89) alleles determine the proinflammatory potential of serum IgA. J Immunol 2007; 178(6): 3973-82.

6. Snir O, Widhe M, von SC, Lindberg J, Padyukov L, Lundberg K et al. Multiple antibody reactivities to citrullinated antigens in sera from patients with rheumatoid arthritis: association with HLA-DRB1 alleles. Ann Rheum Dis 2009; 68(5):736-43.

7. Verheijen R, de Jong BA, van Venrooij WJ. A recombinant topoisomerase I ELISA: screening for IgG, IgM and IgA anti-topo I autoantibodies in human sera. Clin Exp Immunol 1992; 89(3):456-60.

8. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. Arthritis Rheum 1980; 23(5): 581-90.

9. LeRoy EC, Black C, Fleischmajer R, Jablonska S, Krieg T, Medsger TA, Jr. et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. J Rheumatol 1988; 15(2):202-5.

10. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988; 31(3):315-24.

11. Edberg JC, Wu J, Langefeld CD et al. Genetic variation in the CRP promoter: association with systemic lupus erythematosus. Hum Mol Genet. 2008 Apr 15;17(8):1147-55. Epub 2008 Jan 8.

12. Ji JD, Lee WJ, Kong KA, Woo JH, Choi SJ, Lee YH et al. Association of STAT4 polymorphism with rheumatoid arthritis and systemic lupus erythematosus: a meta-analysis. Mol Biol Rep 2010; 37(1):141-7.

13. Rueda B, Broen J, Simeon C, Hesselstrand R, Diaz B, Suarez H et al. The STAT4 gene influences the genetic predisposition to systemic sclerosis phenotype. Hum Mol Genet 2009; 18(11):2071-7.

14. Conley ME, Koopman WJ. Serum IgA1 and IgA2 in normal adults and patients with systemic lupus erythematosus and hepatic disease. Clin Immunol Immunopathol 1983; 26(3):390-7.

15. Svard A, Kastbom A, Reckner-Olsson A, Skogh T. Presence and utility of IgA-class antibodies to cyclic citrullinated peptides in early rheumatoid arthritis: the Swedish TIRA project. Arthritis Res Ther 2008; 10(4):R75.

16. Martínez-Cordero E, Trejo AP, León DE. IgM, IgG, and IgA anti-DNA topoisomerase I antibodies in systemic sclerosis. J Clin Lab Anal. 2009;23(6):408-16

17. Luigi Meroni P, Schur PH, ANA screening: an old test with new recommendations. Ann Rheum Dis 2010;69:1420-1422 doi:10.1136/ard.2009.127100

Chapter 5

A replication study confirms the association of *TNFSF4 (OX40L)* polymorphisms with systemic sclerosis in a large European cohort.

Bossini-Castillo L, Broen JC, Simeon CP, Beretta L, Vonk MC, Ortego-Centeno N, Espinosa G, Carreira P, Camps MT, Navarrete N, González-Escribano MF, Vicente-Rabaneda E, Rodríguez L, Tolosa C, Román-Ivorra JA, Gómez-Gracia I, García-Hernández FJ, Castellví I, Gallego M, Fernández-Nebro A, García-Portales R, Egurbide MV, Fonollosa V, de la Peña PG, Pros A, González-Gay MA, Hesselstrand R, Riemekasten G, Witte T, Coenen MJ, Koeleman BP, Houssiau F, Smith V, de Keyser F, Westhovens R, De Langhe E, Voskuyl AE, Schuerwegh AJ, Chee MM, Madhok R, Shiels P, Fonseca C, Denton C, Claes K, Padykov L, Nordin A, Palm O, Lie BA, Airó P, Scorza R, van Laar JM, Hunzelmann N, Kreuter A, Herrick A, Worthington J, Radstake TR, Martín J*, Rueda B*.

*contributed equally

Annals of the Rheumatic Diseases 2011

Abstract

Objectives: The aim of this study was to confirm the influence of *TNFSF4* polymorphisms on systemic sclerosis (SSc) susceptibility and phenotypic features.

Methods: A total of 8 European populations of Caucasian ancestry were included, comprising 3014 patients with SSc and 3125 healthy controls. Four genetic variants of *TNFSF4* gene promoter (rs1234314, rs844644, rs844648 and rs12039904) were selected as genetic markers.

Results: A pooled analysis revealed the association of rs1234314 and rs12039904 polymorphisms with SSc (OR 1.15, 95% CI 1.02 to 1.31; OR 1.18, 95% CI 1.08 to 1.29, respectively). Significant association of the four tested variants with patients with limited cutaneous SSc (lcSSc) was revealed (rs1234314 OR 1.22, 95% CI 1.07 to 1.38; rs844644 OR 0.91, 95% CI 0.83 to 0.99; rs844648 OR 1.10, 95% CI 1.01 to 1.20 and rs12039904 OR 1.20, 95% CI 1.09 to 1.33). Association of rs1234314, rs844648 and rs12039904 minor alleles with patients positive for anti-centromere antibodies (ACA) remained significant (OR 1.23, 95% CI 1.10 to 1.37; OR 1.12, 95% CI 1.01 to 1.25; OR 1.22, 95% CI 1.07 to 1.38, respectively). Haplotype analysis confirmed a protective haplotype associated with SSc, lcSSc and ACA positive subgroups (OR 0.88, 95% CI 0.82 to 0.96; OR 0.88, 95% CI 0.80 to 0.96; OR 0.86, 95% CI 0.77 to 0.97, respectively) and revealed a new risk haplotype associated with the same groups of patients (OR 1.14, 95% CI 1.03 to 1.26; OR 1.20, 95% CI 1.08 to 1.35; OR 1.23, 95% CI 1.07 to 1.42, respectively).

Conclusions: The data confirm the influence of *TNFSF4* polymorphisms in SSc genetic susceptibility, especially in subsets of patients positive for lcSSc and ACA.

Introduction

Systemic sclerosis (SSc) is a connective tissue disorder characterized by fibrosis, vascular damage and immune imbalance. This pathology has a complex polygenic etiology and variable clinical manifestations. Patients with SSc are commonly classified in two major subgroups: limited cutaneous systemic sclerosis (lcSSc) and diffuse cutaneous systemic sclerosis (dcSSc) (1). Autoantibody status, especially anti-centromere antibodies (ACA) and anti-topoisomerase antibodies (ATA), is clinically used as prognostic bookmaker (1).

Familial clustering and ethnic influences support the genetic component of this disease (2). Initially, only major histocompatibility complex (MHC) genes were firmly associated with SSc. Nevertheless, recently a number of candidate genes such as *STAT4, BANK1 or IRF5*, have been related to SSc genetic predisposition in independent populations by well powered studies (3). Hypothesis free approaches such as genome-wide association studies, have lately confirmed the role of *MHC*, *IRF5* and *STAT4* and uncovered new SSc susceptibility loci, such as *CD247* (4-5).

In this line of research, four *TNFSF4* promoter single nucleotide polymorphisms (SNPs) rs1234314, rs844644, rs844648 and rs12039904 were recently implicated in susceptibility to SSc in a Caucasian American population (composed of 1059 patients with SSc and 698 healthy controls) (6). Interestingly, the *TNFSF4* gene, which encodes OX40L, is considered as a potential autoimmunity candidate gene. OX40L is expressed on activated antigen presenting cells and endothelial cells in acute inflammation. Furthermore, it enhances B cell proliferation and differentiation and its binding to OX40 (CD134) promotes proliferation and survival of T cells (7,8). All these processes could play an important role in loss of immune tolerance and pathology as observed in SSc.

On this basis, the aim of this study was to replicate the association of *TNFSF4* gene promoter polymorphisms with SSc through a large association study in eight independent European populations of Caucasian ancestry, in order to confirm the implication of *TNFSF4* gene in SSc genetic susceptibility and phenotypic features.

Methods

Patients

A total of 3014 cases and 3125 controls from 8 European Caucasian cohorts (Spain, Germany, The Netherlands, Belgium, Italy, Sweden, Norway and UK) were included in this study. Patients with SSc were diagnosed accordingly with the 1980 American College of Rheumatology classification criteria for SSc (9). and were subdivided into those with lcSSc and dcSSc as defined by LeRoy et al. (10).

The following clinical data was collected for ascertainment of clinical phenotype of patients with SSc: age, gender, disease duration and presence of SSc specific autoantibodies, ATA and ACA. Clinical subtype information was available for 82% of the patients, and autoantibody status was available for 74% of the patients. The control population consisted of unrelated healthy individuals recruited in the same geographical regions as patients with SSc, matched by age, sex and ethnicity. The local ethical committees at all participating centres approved the study. Patients and controls were included in the study after written informed consent was obtained.

TNFSF4 polymorphisms genotyping

SNPs rs1234314, rs844644, rs844648 and rs12039904 (tag-SNP of rs2205960 SNP) were genotyped using TaqMan SNP genotyping assays in a 7900HT Real-Time PCR System from Applied Biosystems (Foster City, California, USA). The genotyping call rate was over 93% in all cases and controls included.

Statistical analysis

Association was calculated by 2×2 contingency tables and Fisher's exact test or $\chi 2$ when necessary, obtaining p values, OR and 95% CI using PLINK (V.1.06; http://pngu.mgh.harvard.edu/purcell/plink/). P values below 0.05 after Benjamini and Hochberg False Discovery Rate Method correction were considered as statistically significant. Hardy–Weinberg equilibrium (HWE) was tested for all SNPs at significance level=0.01.

Haplotypes were constructed using PLINK (V.1.06) and Haploview V.4.2 (http://www.broadinstitute.org/haploview/haploview). Haplotypes having a frequency <5% in control groups were excluded for the analysis. Haplotype p values were corrected using Bonferroni correction. Meta-analysis was carried out by PLINK (V. 1.06) and StatsDirect (V.2.6.6; StatsDirect, Altrincham, UK) in the case of haplotypes. Homogeneity among cohorts was calculated using the Breslow–Day method, and OR calculation was performed under fixed effects model (Mantel–Haenszel) or random effects (DerSimonian–Laird) when necessary.

The power of the study for the whole set of patients and controls reached over 98% (Power Calculator for Genetic Studies 2006 (11)).

Results

Analysis of TNFSF4 promoter polymorphisms

The allelic frequencies of the four SNPs tested were similar to those reported for Caucasian populations in previous studies and the international HapMap Project (http://hapmap.ncbi.nlm.nih.gov/) (6,12,13). In addition, the genotypic distribution of healthy controls and SSc cases was in HWE for all SNPs.

Table 1 describes allelic distribution of the four SNPs in the pooled analysis, and supplementary tables 1–4 contain detailed data for each population. Pooled analysis of rs1234314 SNP showed statistically significant association of the G allele with SSc (p =0.03, OR 1.15, 95% CI 1.02 to 1.31), with the subset of patients with lcSSc (p =0.003, OR 1.22, 95% CI 1.07 to 1.38) and with patients positive for ACA (p =2.51E-04, OR 1.23, 95% CI 1.10 to 1.37) (**table 1 and supplementary figure 1**). The association of this genetic marker with lcSSc remained significant after the comparison of this subgroup of patients with those having dcSSc (p =0.01, OR 0.85, 95% CI 0.75 to 0.96, data not shown). Pooled analysis revealed a significant protective association of rs844644 minor allele with lcSSc (p =0.03, OR 0.91, 95% CI 0.83 to 0.99) (**table 1, supplementary figure 2**). Similarly, the rs844648 A allele showed a significant association with susceptibility to lcSSc and ACA positive subgroups (p =0.04, OR 1.10, 95% CI 1.01 to 1.20; p =0.04, OR 1.12, 95% CI 1.01 to 1.25, respectively) (**table 1, supplementary figure 3**). Pooled analysis revealed a strong association of rs12039904 T allele with patients with SSc (p =1.53E-04, OR 1.18, 95% CI 1.08 to 1.29), with patients in the lcSSc subgroup (p =2.81E-04, OR 1.20, 95% CI 1.09 to 1.33) and patients in the ACA-positive subgroup (p =2.09E-03, OR 1.22, 95% CI 1.07 to 1.38) (**table 1, supplementary figure 4**).

TNFSF4 haplotype analysis

Haplotypes represented in over 5% of the healthy controls in any of the eight populations considered, were selected for pooled analysis (**table 2 and supplementary table 5**). Linkage disequilibrium patterns were tested in each of the eight cohorts analysed (supplementary table 6). Only two haplotypes reached significant association with SSc, CAGC (p =2.30E-03, OR 0.88, 95% CI 0.82 to

0.96) and GCAT (*p* =9.10E-03, OR 1.14, 95% CI 1.03 to 1.26) (**supplementary table 5**) (the order of the SNPs is rs1234314-rs844644-rs844648-rs12039904). Interestingly, the CAGC haplotype is composed by the protective alleles of all the tested SNPs while the GCAT haplotype harbours all the risk alleles. The association of CAGC and GCAT haplotypes with SSc clinical features remained significant for patients in the lcSSc (*p* =6.8E-03, OR 0.88, 95% CI 0.80 to 0.96; p=1.3E-03, OR 1.20, 95% CI 1.08 to 1.35, respectively, data not shown) and ACA-positive subsets (*p* =0.01, OR 0.86, 95% CI 0.77 to 0.97; p=3.7E-03, OR 1.23, 95% CI 1.07 to 1.42, respectively, data not shown).

Table 1. Pooled analysis of *TNFSF4* promoter genetic variants

SNP (minor/major alleles), chromosome position (bp)		N	MAF	P_{MH}	P_{FDR}	OR (95% CI)
rs1234314 (G/C), 171444015	**Controls**	2920	0.41			
	SSc	2856	0.44	0.03*†	–	1.15 (1.02 to 1.31)
	lcSSc	1608	0.46	0.003*‡	–	1.22 (1.07 to 1.38)
	dcSSc	724	0.42	0.75§	0.84	1.02 (0.91 to 1.15)
	ACA+	828	0.46	2.51E-04¶	0.001	1.23 (1.10 to 1.37)
	ATA+	519	0.42	0.23**	0.43	1.08 (0.95 to 1.24)
rs844644 (A/C), 171476118	**Controls**	2946	0.47			
	SSc	2912	0.45	0.049†	0.07	0.93 (0.86 to 1.00)
	lcSSc	1653	0.45	0.03‡	0.04	0.91 (0.83 to 0.99)
	dcSSc	743	0.46	0.84§	0.84	0.99 (0.88 to 1.11)
	ACA+	856	0.44	0.049¶	0.049	0.90 (0.80 to 1.00)
	ATA+	533	0.46	0.33**	0.43	0.94 (0.82 to 1.07)
rs844648 (A/G), 171490486	**Controls**	2977	0.43			
	SSc	2940	0.44	0.07†	0.07	1.07 (1.00 to 1.15)
	lcSSc	1673	0.45	0.04‡	0.04	1.1 (1.01 to 1.20)
	dcSSc	742	0.42	0.69§	0.84	0.98 (0.87 to 1.10)
	ACA+	860	0.45	0.04¶	0.049	1.12 (1.01 to 1.25)
	ATA+	529	0.42	0.74**	0.74	1.02 (0.89 to 1.17)

rs12039904 (T/C), 171478896	**Controls**	2991	0.23			
	SSc	2894	0.26	1.53E-04†	6.12E-04	1.18 (1.08 to 1.29)
	lcSSc	1639	0.26	2.81E-04‡	5.61E-04	1.20 (1.09 to 1.33)
	dcSSc	735	0.24	0.3§	0.84	1.07 (0.94 to 1.23)
	ACA+	840	0.26	2.09E-03¶	4.17E-03	1.22 (1.07 to 1.38)
	ATA+	523	0.24	0.15**	0.43	1.12 (0.96 to 1.31)

Controls are used as reference for all comparisons. * DerSimonian–Laird random effects model p value. † Breslow–Day rs1234314 p=0.01; rs844644 p=0.23; rs844648 p=0.52; rs12039904 p=0.50. ‡ Breslow–Day rs1234314 p=0.08; rs844644 p=0.52; rs844648 p=0.33; rs12039904 p=0.48. § Breslow–Day rs1234314 p=0.61; rs844644 p=0.49; rs844648 p=1.00; rs12039904 p=0.38. ¶ Breslow–Day rs1234314 p=0.29; rs844644 p=0.79; rs844648 p=0.94; rs12039904 p=0.74. ** Breslow–Day rs1234314 p=0.41; rs844644 p=0.73; rs844648 p=0.71; rs12039904 p=0.56.

Table 2. Pooled analysis of *TNFSF4* haplotypes in patients with systemic sclerosis and controls

(2n cases/2n controls)	Haplotype	Cases (%)	Controls (%)	P_{MH}	OR (95% CI)
Pooled (5222/5296)	**CAGC**	40.49	43.19	2.30E-03*	0.88 (0.82 to 0.96)
	CCGC	10.23	10.04	0.47†‡	1.09 (0.86 to 1.39)
	GAGC	4.43	3.75	0.08§	1.20 (0.98 to 1.47)
	GCAC	16.71	16.68	0.53¶	0.97 (0.87 to 1.07)
	GCAT	21.54	19.33	9.10E-03**	1.14 (1.03 to 1.26)
	Others	6.60	7.01	0.67*††	0.94 (0.73 to 1.23)

The order of the SNPs is rs1234314-rs844644-rs844648-rs12039904.* Breslow–Day p=0.22. † DerSimonian–Laird random effects model p value. ‡ Breslow–Day p=0.004. § Breslow–Day p=0.62. ¶ Breslow–Day p=0.16. ** Breslow–Day p=0.24. †† Breslow–Day p=0.01. PMH, allelic Mantel–Haenszel fixed effects model p value.

Discussion

TNFSF4 polymorphisms have been related to susceptibility for different autoimmune diseases including SSc. (6, 14) With the aim of validating the initially reported association of *TNFSF4* gene in SSc, we conducted a large case-control study and a pooled analysis in eight independent European populations of Caucasian ancestry.

In accordance with the report by Gourh et al, our study supports the implication of *TNFSF4* gene promoter polymorphisms in SSc genetic predisposition (6). Stratification by SSc clinical subtype or autoantibody status confirmed the significant association of the *TNFSF4* variants with the patients in the lcSSc subset and ACA-positive subgroup but not with patients in the dcSSc or ATA-positive subsets. Nevertheless, the risk or protective directions in the associations were consistent with those reported by Gourh et al. (6).

Similarly, haplotype pooled analysis results obtained in the present study keep in with the findings from Gourh et al. (6). The most represented haplotype in both reports are equivalent, appear in similar frequency and have a protective effect. Nevertheless, in our study the opposite haplotype GCAT, which could not be observed in the previous study, showed a significant risk association with SSc.

Interestingly, previous findings in systemic lupus erythematosus revealed the existence of equivalent protective and risk haplotypes to the ones reported in this study. Moreover, functional data showed that the risk haplotype produced an increased level of *TNFSF4* transcript (compared to the protective haplotype), and a higher surface expression of OX40L in lymphoblastoid cell lines and peripheral blood lymphocytes after activation. This overexpression seems to be related to the destruction of the DNA binding site for the transcriptional repressor E4BP4 (with a role in the survival of early B cell progenitors) (12). Thus, the *TNFSF4* risk haplotype associated with SSc and producing higher levels of OX40L might be implicated in the pathogenic mechanisms of SSc, by the alteration of regulatory processes controlling B and T cell proliferation and differentiation, leading to autoantibody production and tissue damage (7 8 14–16). Further studies are necessary to elucidate the exact molecular mechanisms by which OX40L is implicated in SSc pathogenesis and more precisely how it can lead to the development of lcSSc and ACA production.

In summary, our results confirm the implication of *TNFSF4* promoter polymorphisms in SSc susceptibility, especially in patients in the lcSSc and ACA-positive subgroups. These findings together with previous genetic and functional studies suggest *TNFSF4* as an interesting and consistent genetic factor for SSc and other autoimmune diseases and may open new opportunities for SSc treatment.

References

1. Steen VD. The many faces of scleroderma. Rheum Dis Clin North Am 2008;34:1–15; v.

2. Agarwal SK, Tan FK, Arnett FC. Genetics and genomic studies in scleroderma (systemic sclerosis). Rheum Dis Clin North Am 2008;34:17–40; v.

3. Agarwal SK, Reveille JD. The genetics of scleroderma (systemic sclerosis). Curr Opin Rheumatol 2010;22:133–8.

4. Zhou X, Lee JE, Arnett FC, et al. HLA-DPB1 and DPB2 are genetic loci for systemic sclerosis: a genome-wide association study in Koreans with replication in North Americans. Arthritis Rheum 2009;60:3807–14.

5. Radstake TR, Gorlova O, Rueda B, et al. Genome-wide association study of systemic sclerosis identifies CD247 as a new susceptibility locus. Nat Genet 2010;42:426–9.

6. Gourh P, Arnett FC, Tan FK, et al. Association of TNFSF4 (OX40L) polymorphisms with susceptibility to systemic sclerosis. Ann Rheum Dis 2010;69:550–5.

7. Manku H, Graham DS, Vyse TJ. Association of the co-stimulator OX40L with systemic lupus erythematosus. J Mol Med 2009;87:229–34.

8. Gough MJ, Weinberg AD. OX40 (CD134) and OX40L. Adv Exp Med Biol 2009;647:94–107.

9. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. Arthritis Rheum 1980;23:581–90.

10. LeRoy EC, Black C, Fleischmajer R, et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. J Rheumatol 1988;15:202–5.

11. Skol AD, Scott LJ, Abecasis GR, et al. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. Nat Genet 2006;38:209–13.

12. Cunninghame Graham DS, Graham RR, Manku H, et al. Polymorphism at the TNF superfamily gene TNFSF4 confers susceptibility to systemic lupus erythematosus. Nat Genet 2008;40:83–9.

13. Delgado-Vega AM, Abelson AK, Sánchez E, et al. Replication of the TNFSF4 (OX40L) promoter region association with systemic lupus erythematosus. Genes Immun 2009;10:248–53.

14. Croft M. Control of immunity by the TNFR-related molecule OX40 (CD134). Annu Rev Immunol 2010;28:57–78.

15. Croft M, So T, Duan W, et al. The significance of OX40 and OX40L to T-cell biology and immune disease. Immunol Rev 2009;229:173–91.

16. Radstake TR, van Bon L, Broen J, et al. Increased frequency and compromised function of T regulatory cells in systemic sclerosis (SSc) is related to a diminished CD69 and TGFbeta expression. PLoS ONE 2009;4:e5981.

Supplementary data

Supplementary Table 1. Distribution of *TNFSF4* rs1234314 genetic variant in eight replication cohorts.

Population		N	MAF	P_u	P_{FDR}	OR (95% CI)
Spain	**Controls**	905	0.44			
	SSc	853	0.44	0.91	0.91	0.99 (0.87-1.13)
	lcSSc	491	0.45	0.53	0.64	1.05 (0.90-1.23)
	dcSSc	236	0.42	0.53	0.98	0.94 (0.76-1.15)
	ACA+	327	0.45	0.57	0.57	1.05 (0.88-1.26)
	ATA+	169	0.41	0.33	0.66	0.89(0.70-1.13)
Germany	**Controls**	245	0.41			
	SSc	308	0.46	0.26	0.53	1.15 (0.90-1.46)
	lcSSc	146	0.46	0.14	0.31	1.25 (0.93-1.62)
	dcSSc	126	0.45	0.29	0.56	1.18 (0.87-1.60)
	ACA+	73	0.45	0.43	0.94	1.16 (0.80-1.69)
	ATA+	46	0.48	0.21	0.79	1.33 (0.85-2.08)
The Netherlands	**Controls**	187	0.41			
	SSc	329	0.49	0.03	0.1	1.34 (1.04-1.73)
	lcSSc	101	0.5	0.06	0.13	1.39 (0.98-1.95)
	dcSSc	30	0.38	0.65	0.99	0.88 (0.50-1.54)
	ACA+	35	0.54	0.05	0.12	1.68 (1.00-2.80)
	ATA+	27	0.48	0.35	0.62	1.31 (0.74-2.33)
Belgium	**Controls**	251	0.42			
	SSc	174	0.5	0.02	0.03	1.37 (1.04-1.81)
	lcSSc	111	0.53	0.01	0.02	1.51 (1.10-2.08)
	dcSSc	52	0.44	0.74	0.74	1.08 (0.7-1.65)
	ACA+	38	0.51	0.15	0.56	1.43 (0.88-2.32)
	ATA+	29	0.55	0.06	0.09	1.67 (0.97-2.89)
Italy	**Controls**	465	0.36			
	SSc	483	0.42	0.02	0.07	1.25 (1.04-1.50)
	lcSSc	302	0.43	0.01	0.04	1.31 (1.07-1.62)
	dcSSc	103	0.35	0.83	0.9	0.97 (0.71-1.32)
	ACA+	159	0.46	2.92E-03	0.01	1.48 (1.14-1.91)
	ATA+	171	0.4	0.25	0.49	1.16 (0.90-1.50)
Sweden	**Controls**	246	0.41			
	SSc	179	0.36	0.13	0.53	0.81 (0.61-1.07)

Population		N	MAF	P_u	P_{FDR}	OR (95% CI)
	lcSSc	78	0.37	0.34	0.66	0.83 (0.57-1.21)
	dcSSc	36	0.33	0.22	0.59	0.72 (0.43-1.22)
	ACA+	27	0.39	0.78	0.98	0.92 (0.52-1.64)
	ATA+	19	0.37	0.63	0.77	0.84 (0.43-1.67)
Norway	**Controls**	264	0.41			
	SSc	83	0.42	0.81	0.81	1.04 (0.73-1.49)
	lcSSc	49	0.41	0.96	0.96	1.01 (0.65-1.57)
	dcSSc	28	0.45	0.55	0.97	1.18 (0.68-2.06)
	ACA+	40	0.45	0.45	0.45	1.20 (0.75-1.93)
	ATA+	11	0.41	0.97	0.97	1.02 (0.43-2.42)
United Kingdom	**Controls**	357	0.4			
	SSc	447	0.47	0.002	0.008	1.37 (1.12-1.67)
	lcSSc	330	0.48	1.72E-03	0.007	1.41 (1.14-1.74)
	dcSSc	113	0.45	0.18	0.7	1.23 (0.91-1.67)
	ACA+	129	0.48	0.02	0.07	1.41 (1.06-1.88)
	ATA+	47	0.43	0.59	0.9	1.13 (0.73-1.74)

Nucleotide change: G/C. Position in chromosome 1: 171444015bp. Controls are used as reference for all comparisons. MAF: Minor allele (G) frequency; P_u: Allelic chi-square uncorrected p-value; P_{FDR}: Corrected p-value using Benjamini & Hochberg False Discovery Rate

Supplementary Table 2. Distribution of *TNFSF4* rs844644 genetic variant in eight replication cohorts.

Population		N	MAF	Pu	P_{FDR}	OR (95% CI)
Spain	**Controls**	902	0.45			
	SSc	880	0.44	0.56	0.91	0.96 (0.84-1.10)
	lcSSc	511	0.43	0.32	0.64	0.93 (0.79-1.08)
	dcSSc	240	0.45	0.98	0.98	1 (0.81-1.22)
	ACA+	335	0.42	0.16	0.33	0.88 (0.74-1.05)
	ATA+	171	0.46	0.7	0.73	1.05 (0.83-1.32)
Germany	**Controls**	265	0.42			
	SSc	341	0.46	0.18	0.53	1.17 (0.93-1.47)
	lcSSc	159	0.47	0.15	0.31	1.22 (0.93-1.62)
	dcSSc	141	0.45	0.42	0.56	1.13 (0.84-1.51)
	ACA+	81	0.46	0.47	0.94	1.14 (0.80-1.62)
	ATA+	52	0.43	0.88	0.88	1.03 (0.68-1.58)
The Netherlands	**Controls**	171	0.48			
	SSc	326	0.43	0.17	0.17	0.83 (0.64-1.08)
	lcSSc	96	0.46	0.64	0.64	0.92 (0.64-1.31)
	dcSSc	29	0.45	0.66	0.99	0.88 (0.50-1.54)
	ACA+	32	0.41	0.28	0.28	0.74 (0.43-1.28)
	ATA+	26	0.44	0.62	0.62	0.86 (0.48-1.55)
Belgium	**Controls**	238	0.45			
	SSc	170	0.38	0.06	0.06	0.76 (0.58-1.01)
	lcSSc	109	0.4	0.21	0.22	0.81 (0.59-1.13)
	dcSSc	51	0.35	0.06	0.24	0.66 (0.42-1.02)
	ACA+	37	0.46	0.93	0.93	1.02 (0.63-1.67)
	ATA+	30	0.33	0.08	0.09	0.6 (0.34-1.06)
Italy	**Controls**	478	0.51			
	SSc	469	0.47	0.11	0.21	0.86 (0.72-1.03)
	lcSSc	306	0.47	0.13	0.25	0.85 (0.70-1.05)
	dcSSc	102	0.52	0.71	0.9	1.06 (0.78-1.43)
	ACA+	167	0.47	0.27	0.31	0.87 (0.68-1.12)
	ATA+	169	0.46	0.2	0.49	0.85 (0.66-1.09)
Sweden	**Controls**	269	0.48			
	SSc	179	0.5	0.53	0.73	1.09 (0.83-1.42)
	lcSSc	80	0.46	0.66	0.66	0.92 (0.65-1.32)
	dcSSc	36	0.54	0.29	0.59	1.3 (0.79-2.13)

Population		N	MAF	Pu	P_{FDR}	OR (95% CI)
	ACA+	28	0.46	0.87	0.98	0.95 (0.55-1.66)
	ATA+	20	0.5	0.77	0.77	1.1 (0.58-2.09)
Norway	**Controls**	265	0.48			
	SSc	94	0.45	0.42	0.7	0.87 (0.62-1.22)
	lcSSc	58	0.41	0.19	0.37	0.76 (0.51-1.14)
	dcSSc	30	0.5	0.78	0.97	1.08 (0.63-1.84)
	ACA+	46	0.39	0.11	0.24	0.69 (0.44-1.09)
	ATA+	13	0.46	0.85	0.97	0.92 (0.42-2.04)
United Kingdom	**Controls**	358	0.49			
	SSc	453	0.45	0.16	0.32	0.87 (0.71-1.06)
	lcSSc	334	0.45	0.21	0.38	0.87 (0.71-1.08)
	dcSSc	114	0.45	0.35	0.7	0.87 (0.64-1.17)
	ACA+	130	0.47	0.54	0.72	0.92 (0.69-1.22)
	ATA+	52	0.48	0.9	0.9	0.97 (0.65-1.47)

Nucleotide change: A/C. Position in chromosome 1: 171476118bp. Controls are used as reference for all comparisons. MAF: Minor allele (A) frequency; P_u: Allelic chi-square uncorrected p-value; P_{FDR}: Corrected p-value using Benjamini & Hochberg False Discovery Rate

Supplementary Table 3. Distribution of *TNFSF4* rs844648 genetic variant in eight replication cohorts.

Population		N	MAF	Pu	P_{FDR}	OR (95% CI)
Spain	**Controls**	865	0.45			
	SSc	887	0.44	0.8	0.91	1.02 (0.87-1.16)
	lcSSc	510	0.45	0.64	0.64	1.04 (0.89-1.21)
	dcSSc	240	0.44	0.77	0.98	0.97 (0.79-1.19)
	ACA+	336	0.46	0.38	0.5	1.08 (0.91-1.30)
	ATA+	171	0.43	0.73	0.73	0.96 (0.76-1.21)
Germany	**Controls**	271	0.44			
	SSc	342	0.43	0.82	0.82	0.97 (0.78-1.22)
	lcSSc	160	0.45	0.82	0.82	1.03 (0.78-1.36)
	dcSSc	140	0.42	0.56	0.56	0.92 (0.68-1.23)
	ACA+	77	0.44	0.96	0.98	1.01 (0.70-1.45)
	ATA+	50	0.41	0.59	0.79	0.89 (0.58-1.37)
The Netherlands	**Controls**	199	0.43			
	SSc	327	0.48	0.13	0.17	1.22 (0.95-1.56)
	lcSSc	104	0.48	0.28	0.37	1.20 (0.86-1.69)
	dcSSc	31	0.44	0.99	0.99	1.00 (0.58-1.72)
	ACA+	35	0.51	0.22	0.28	1.38 (0.83-2.29)
	ATA+	28	0.48	0.5	0.62	1.21 (0.69-2.12)
Belgium	**Controls**	278	0.43			
	SSc	176	0.5	0.02	0.03	1.37 (1.04-1.79)
	lcSSc	113	0.53	0.01	0.02	1.5 (1.10-2.06)
	dcSSc	52	0.45	0.61	0.74	1.12 (0.73-1.70)
	ACA+	38	0.47	0.42	0.56	1.22 (0.75-1.97)
	ATA+	29	0.55	0.06	0.09	1.67 (0.97-2.87)
Italy	**Controls**	469	0.37			
	SSc	479	0.39	0.36	0.36	1.09 (0.91-1.31)
	lcSSc	311	0.4	0.27	0.36	1.13 (0.91-1.39)
	dcSSc	106	0.36	0.9	0.9	0.98 (0.72-1.34)
	ACA+	169	0.41	0.19	0.31	1.19 (0.92-1.53)
	ATA+	172	0.37	0.96	0.96	1.01 (0.78-1.30)
Sweden	**Controls**	266	0.43			
	SSc	185	0.42	0.73	0.73	0.95 (0.73-1.25)
	lcSSc	79	0.38	0.26	0.66	0.81 (0.56-1.17)
	dcSSc	36	0.42	0.82	0.93	0.95 (0.57-1.56)

Population		N	MAF	Pu	P_{FDR}	OR (95% CI)
	ACA+	27	0.43	0.95	0.98	0.98 (0.56-1.73)
	ATA+	20	0.4	0.71	0.77	0.88 (0.46-1.70)
Norway	**Controls**	277	0.44			
	SSc	94	0.46	0.52	0.7	1.12 (0.80-1.56)
	lcSSc	58	0.49	0.27	0.37	1.25 (0.84-1.87)
	dcSSc	30	0.43	0.97	0.97	0.99 (0.58-1.70)
	ACA+	46	0.51	0.18	0.24	1.35 (0.87-2.10)
	ATA+	13	0.42	0.9	0.97	0.95 (0.43-2.10)
United Kingdom	**Controls**	352	0.43			
	SSc	450	0.45	0.45	0.45	1.08 (0.88-1.32)
	lcSSc	338	0.45	0.38	0.38	1.1 (0.89-1.36)
	dcSSc	107	0.43	0.92	0.92	0.98 (0.72-1.34)
	ACA+	132	0.45	0.61	0.72	1.08 (0.81-1.43)
	ATA+	46	0.46	0.62	0.9	1.12 (0.72-1.73)

Nucleotide change: A/G. Position in chromosome 1: 171490486bp. Controls are used as reference for all comparisons. MAF: Minor allele frequency; P_u: Allelic chi-square uncorrected p-value; P_{FDR}: Corrected p-value using Benjamini & Hochberg False Discovery Rate

Supplementary Table 4. Distribution of *TNFSF4* rs12039904 genetic variant in eight replication cohorts.

Population		N	MAF	Pu	P_{FDR}	OR (95% CI)
Spain	**Controls**	902	0.22			
	SSc	863	0.26	0.01	0.06	1.21 (1.04-1.42)
	lcSSc	499	0.26	0.01	0.05	1.25 (1.05-1.50)
	dcSSc	238	0.25	0.22	0.86	1.16 (0.92-1.47)
	ACA+	329	0.27	0.02	0.07	1.29 (1.05-1.58)
	ATA+	171	0.26	0.09	0.36	1.26 (0.96-1.64)
Germany	**Controls**	272	0.26			
	SSc	312	0.25	0.67	0.82	0.94 (0.73-1.23)
	lcSSc	149	0.28	0.54	0.73	1.10 (0.80-1.52)
	dcSSc	132	0.2	0.09	0.35	0.74 (0.52-1.05)
	ACA+	73	0.26	0.98	0.98	1.01 (0.66-1.53)
	ATA+	45	0.22	0.46	0.79	0.82 (0.48-1.39)
The Netherlands	**Controls**	167	0.19			
	SSc	338	0.26	0.06	0.11	1.36 (0.99-1.86)
	lcSSc	102	0.28	0.048	0.13	1.50 (1.00-2.24)
	dcSSc	31	0.19	0.77	0.99	0.91 (0.46-1.79)
	ACA+	35	0.31	0.06	0.12	1.72 (0.98-3.05)
	ATA+	27	0.28	0.26	0.62	1.45 (0.76-2.78)
Belgium	**Controls**	262	0.23			
	SSc	175	0.31	0.009	0.03	1.5 (1.10-2.04)
	lcSSc	112	0.31	0.01	0.02	1.55 (1.09-2.19)
	dcSSc	52	0.3	0.12	0.24	1.45 (0.91-2.31)
	ACA+	37	0.28	0.28	0.56	1.35 (0.78-2.33)
	ATA+	29	0.33	0.08	0.09	1.66 (0.92-2.98)
Italy	**Controls**	493	0.21			
	SSc	473	0.23	0.17	0.23	1.16 (0.94-1.44)
	lcSSc	307	0.22	0.51	0.51	1.09 (0.85-1.39)
	dcSSc	105	0.24	0.23	0.9	1.24 (0.87-1.76)
	ACA+	168	0.23	0.31	0.31	1.17 (0.87-1.57)
	ATA+	168	0.22	0.66	0.88	1.07 (0.79-1.45)
Sweden	**Controls**	264	0.26			
	SSc	186	0.28	0.54	0.73	1.1 (0.81-1.48)
	lcSSc	76	0.24	0.66	0.66	0.91 (0.60-1.38)
	dcSSc	37	0.26	0.93	0.93	0.98 (0.56-1.70)

Population		N	MAF	Pu	P_{FDR}	OR (95% CI)
	ACA+	25	0.26	0.98	0.98	0.99 (0.51-1.92)
	ATA+	20	0.23	0.61	0.77	0.82 (0.38-1.77)
Norway	**Controls**	274	0.25			
	SSc	93	0.28	0.43	0.7	1.16 (0.80-1.69)
	lcSSc	58	0.31	0.18	0.37	1.35 (0.87-2.09)
	dcSSc	27	0.24	0.89	0.97	0.95 (0.51-1.80)
	ACA+	46	0.33	0.12	0.24	1.45 (0.90-2.34)
	ATA+	13	0.23	0.82	0.97	0.9 (0.35-2.29)
United Kingdom	**Controls**	357	0.24			
	SSc	454	0.27	0.26	0.35	1.14 (0.91-1.43)
	lcSSc	336	0.26	0.3	0.38	1.14 (0.89,1.45)
	dcSSc	113	0.26	0.54	0.72	1.11 (0.79-1.57)
	ACA+	127	0.25	0.72	0.72	1.06 (0.76-1.48)
	ATA+	50	0.23	0.81	0.9	0.94 (0.57-1.55)

Nucleotide change: T/C. Position in chromosome 1: 171478896bp. Controls are used as reference for all comparisons. MAF: Minor allele frequency; P_u: Allelic chi-square uncorrected p-value; P_{FDR}: Corrected p-value using Benjamini & Hochberg False Discovery Rate

Supplementary Table 5. Distribution of *TNFSF4* haplotypes in SSc patients and controls in eight replication cohorts.

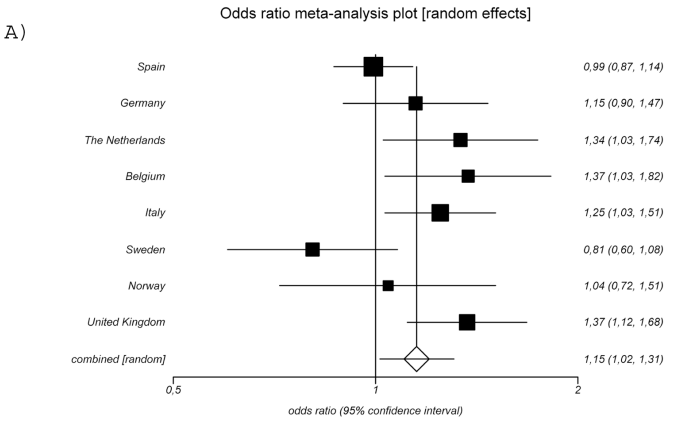
Population (2n cases/2n controls)		Cases (%)	Controls (%)	P_{Bf}	OR (95% CI)
Spain					
(1650/1666)	**CAGC**	40.77	41.43	NS	0.97 (0.85-1.12)
	CCGC	9.66	11.37	0.55	0.83 (0.67-1.04)
	GAGC	3.14	2.77	NS	1.14 (0.76-1.70)
	GCAC	17.16	19.42	NS	0.98 (0.88-1.08)
	GCAT	21.59	19.67	0.88	1.12 (0.95-1.33)
	Others	7.69	5.35	0.03	1.47 (1.11-1.95)
Germany	**CAGC**	41.80	39.30	NS	1.11 (0.86-1.43)
(534/468)	**CCGC**	9.96	13.32	0.51	0.72 (0.49-1.07)
	GAGC	5.86	3.28	0.28	1.78 (0.95-3.33)
	GCAC	17.77	16.16	NS	1.02 (0.85-1.22)
	GCAT	19.53	20.96	NS	0.92 (0.67-1.25)
	Others	5.08	6.99	NS	0.72 (0.42-1.22)
The Netherlands	**CAGC**	39.01	45.56	0.40	0.76 (0.57-1.03)
(574/248)	**CCGC**	6.56	6.05	NS	1.05 (0.57-1.94)
	GAGC	4.79	3.63	NS	1.25 (0.59-2.65)
	GCAC	19.15	18.95	NS	1.01 (0.82-1.26)
	GCAT	22.52	16.53	0.26	1.45 (0.98-2.13)
	Others	7.98	9.27	NS	0.83 (0.49-1.40)
Belgium	**CAGC**	34.89	40.23	0.68	0.8 (0.59-1.08)
(328/438)	**CCGC**	9.03	9.77	NS	0.93 (0.57-1.52)
	GAGC	4.05	5.35	NS	0.77 (0.39-1.53)
	GCAC	19.63	16.51	NS	1.01 (0.82-1.24)
	GCAT	25.23	20.23	0.52	1.33 (0.94-1.88)
	Others	7.17	7.91	NS	0.91 (0.53-1.57)
Italy	**CAGC**	41.57	48.14	0.03	0.77 (0.63-0.93)
(866/878)	**CCGC**	13.00	10.44	0.50	1.28 (0.95-1.72)
	GAGC	5.27	3.83	0.76	1.39 (0.88-2.19)
	GCAC	14.87	14.04	NS	1.02 (0.89-1.17)
	GCAT	20.14	16.71	0.33	1.26 (0.98-1.60)
	Others	5.15	6.84	0.70	0.74 (0.5-1.11)
Sweden	**CAGC**	39.27	42.40	NS	0.88 (0.65-1.18)
(310/450)	**CCGC**	18.48	9.52	1.93E-03	2.14 (1.39-3.28)
	GAGC	3.30	4.08	NS	0.83 (0.39-1.81)

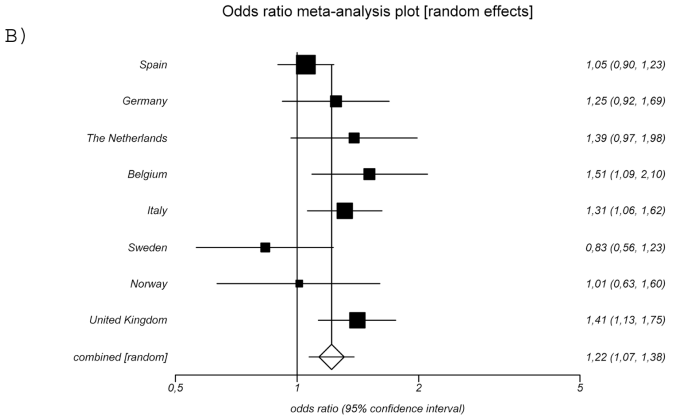
Population (2n cases/2n controls)		Cases (%)	Controls (%)	P_{Bf}	OR (95% CI)
	GCAC	8.25	13.83	NS	0.95 (0.77-1.18)
	GCAT	17.16	20.86	NS	0.79 (0.54-1.15)
	Others	13.53	9.30	0.35	1.52 (0.96-2.41)
Norway	**CAGC**	40.49	42.37	NS	0.93 (0.65-1.33)
(164/472)	**CCGC**	10.43	8.39	NS	1.31 (0.72-2.37)
	GAGC	5.52	5.59	NS	1.05 (0.49-2.26)
	GCAC	10.43	13.12	NS	1.01 (0.78-1.31)
	GCAT	25.15	21.29	NS	1.25 (0.83-1.9)
	Others	7.98	9.25	NS	0.89 (0.47-1.69)
United Kingdom	**CAGC**	40.29	45.18	0.31	0.82 (0.66-1.01)
(796/676)	**CCGC**	8.74	7.83	NS	1.12 (0.77-1.64)
	GAGC	5.35	4.07	NS	1.32 (0.80-2.16)
	GCAC	18.38	16.42	NS	1.02 (0.88-1.19)
	GCAT	22.03	19.28	1.00	1.18 (0.91-1.53)
	Others	5.22	7.23	0.57	0.71 (0.46-1.09)

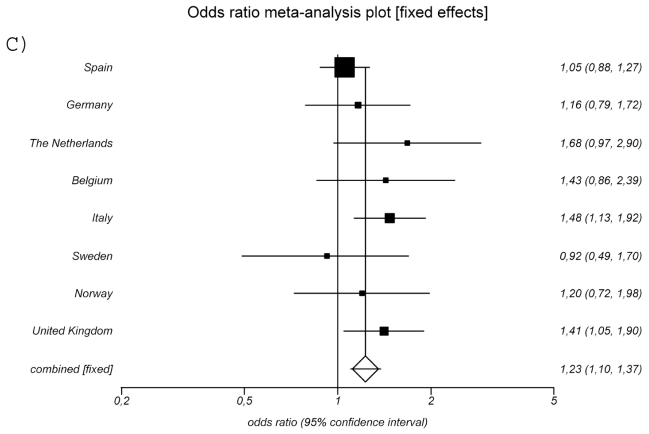
The order of the SNPs is rs1234314-rs844644-rs844648-rs12039904. $P_{Bf:}$ Corrected p-value using Bonferroni multiple test correction; NS: non-significant; OR: odds ratio.

Supplementary Table 6. Linkage disequilibrium of *TNFSF4* polymorphisms in eight European cohorts.

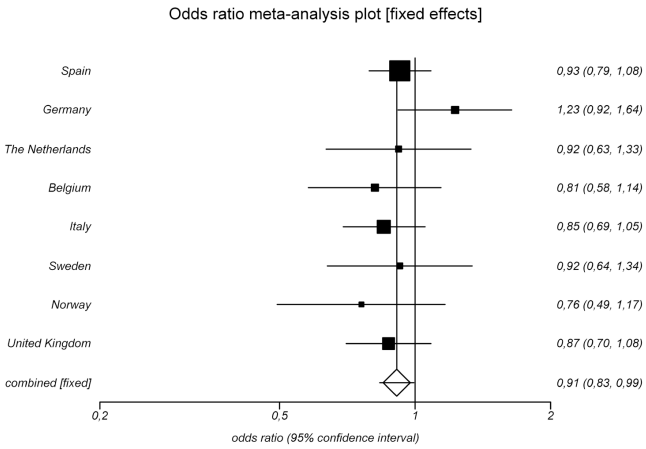
Population	R^2 rs1234314-rs844644	R^2 rs1234314-rs844648	R^2 rs1234314-rs12039904	R^2 rs844644-rs844648	R^2 rs844644-rs12039904	R^2 rs844648-rs12039904
Spain	0.44	0.68	0.24	0.55	0.24	0.33
Germany	0.31	0.63	0.22	0.51	0.28	0.42
The Netherlands	0.48	0.56	0.13	0.72	0.23	0.29
Belgium	0.38	0.52	0.27	0.61	0.28	0.36
Italy	0.39	0.55	0.26	0.58	0.25	0.35
Sweden	0.37	0.45	0.22	0.66	0.39	0.41
United Kingdom	0.41	0.58	0.18	0.68	0.27	0.34
Norway	0.31	0.44	0.22	0.66	0.27	0.35



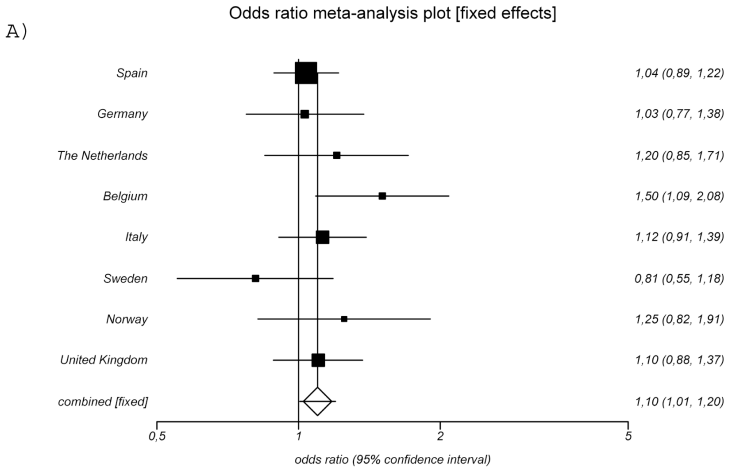


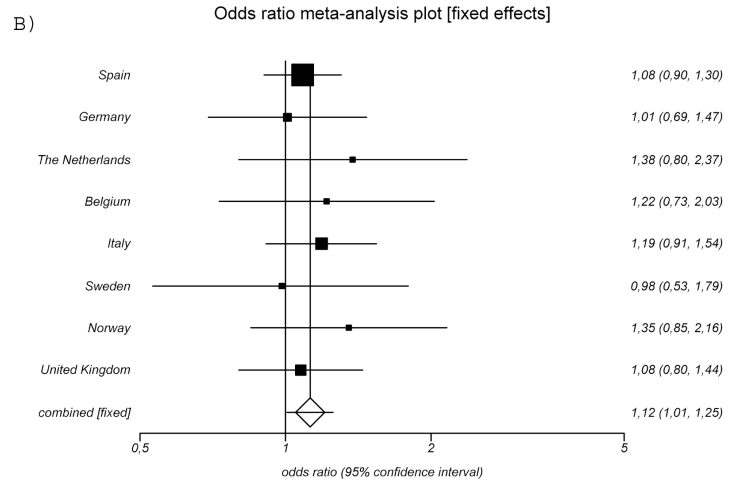


Supplementary figure 1.A) Forest plot of the *TNFSF4* rs1234314 polymorphism in SSc versus controls in eight cohorts, B) in lcSSc patients versus controls and C) ACA positive patients versus controls.

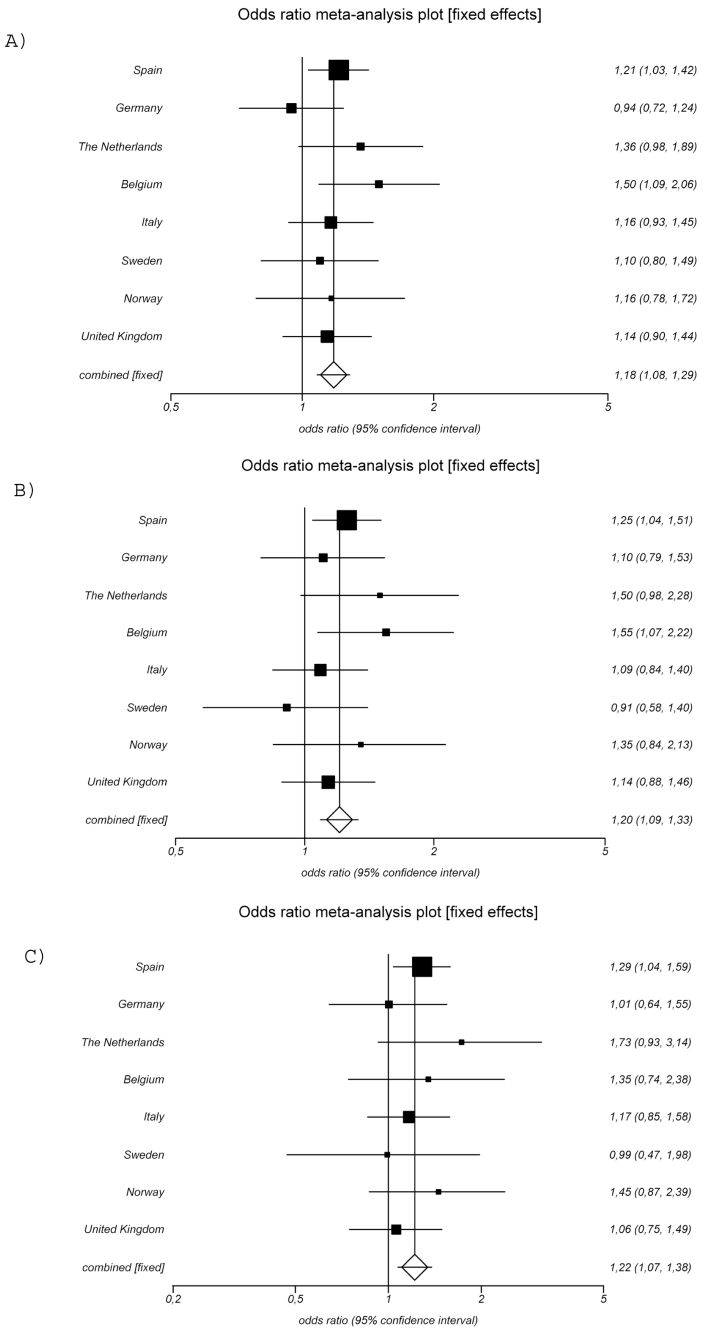


Supplementary figure 2. Forest plot for the meta-analysis of the *TNFSF4* rs844644 polymorphism in lcSSc patients versus controls in eight Caucasian cohorts.

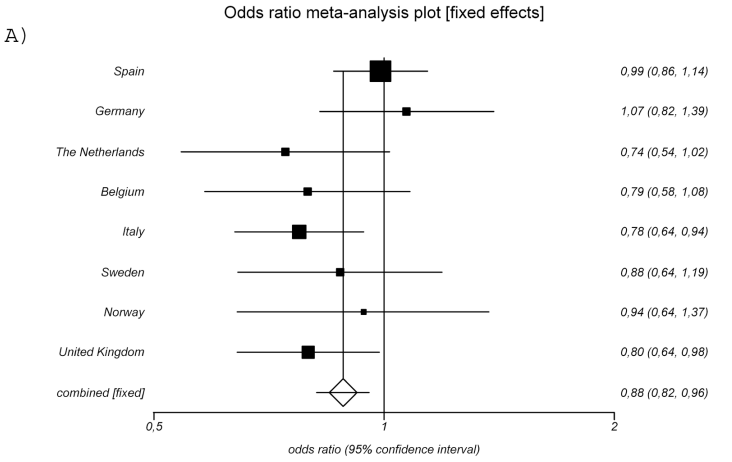


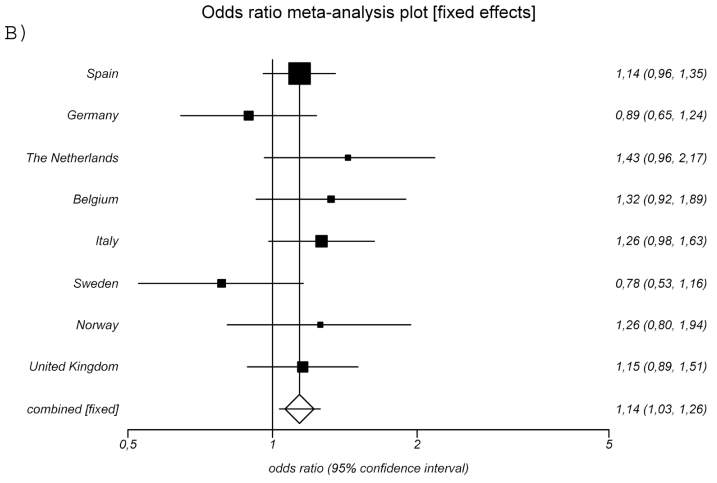


Supplementary Figure 3.A) Forest plot for the meta-analysis of the *TNFSF4* rs844648 polymorphism in lcSSc patients versus controls in eight Caucasian cohorts. B) Forest plot for the meta-analysis of the *TNFSF4* rs844648 polymorphism in ACA positive patients versus controls in eight Caucasian cohorts.



Supplementary figure 4.A) Forest plot of the *TNFSF4* rs12039904 polymorphism in lSSc patients versus controls B) lcSSc patients versus controls and C) ACA positive patients versus controls.





Supplementary figure 5.A) Forest plot for the meta-analysis of the in CAGC haplotype in SSc patients versus controls in eight Caucasian cohorts. B) Forest plot for the meta-analysis of GCAT haplotype in SSc patients versus controls in eight Caucasian cohorts.

Part II

Genetic association studies with functional and clinical validation

Paul Klee, *Comedian's Handbill*, 1938

Chapter 6

Polymorphisms in the *Interleukin 4, Interleukin 13* and corresponding receptor genes are not associated with Systemic Sclerosis and do not influence gene expression.

Broen JCA, Dieude P, Vonk MC, Beretta L, Carmona FD, Herrick A, Worthington J, Hunzelmann N, Riemekasten G, Kiener H, Scorza R, Simeon CP, Fonollosa V (for the Spanish Systemic Sclerosis group), Carreira P, Ortego-Centeno N, Gonzalez-Gay MA, Airo' P, Coenen MJH, Tsang K, Aliprantis AO, Martin J*, Allanore Y*, and Radstake TRDJ*

The Journal of Rheumatology February 2012

Abstract

Aim: Polymorphisms in the genes encoding *interleukin 4 (IL4)*, *interleukin 13 (IL13)* and their corresponding receptors have previously been associated with multiple immune mediated diseases. In this study we aim to validate these previous observations in systemic sclerosis (SSc) patients and scrutinize the effect of the polymorphisms on gene expression in various populations of peripheral blood leukocytes.

Patients and Methods: We genotyped a cohort of 2488 SSc patients and 2246 healthy controls from The Netherlands, Spain, United Kingdom, Italy, Germany and France. Taqman assays were used to genotype single nucleotide polymorphisms (SNP)s in the following genes: 1) *IL4* (*−590C>T*/rs2243250), 2) *IL4 receptor alpha (IL4RA)* (*Q576R*/rs1801275) 3) *IL13* (*R130Q*/rs20541 and *-1112C>T* rs1800925) and 4) *IL13 receptor alpha 1* (*IL13RA1)* (*43163G>A/rs6646259).* In addition, the effect of these polymorphisms on expression of the corresponding genes was assessed using quantitative RT-PCR on RNA derived from peripheral blood B-cells, T-cells, plasmacytoid dendritic cells, monocytes and myeloid dendritic cells. Moreover, we investigated whether these polymorphisms influence development of pulmonary complications in SSc patients over 15 years.

Results: None of the investigated polymorphisms were associated with SSc or any SSc clinical subtype. In addition, we did not observe any effect on transcript levels in the cell subtypes or on development of pulmonary complications.

Conclusions: Our data show that several polymorphisms in *IL4, IL13* and their receptors do not play a role in SSc and do not influence the expression of their corresponding transcript in peripheral blood cells.

Introduction

Systemic Sclerosis (SSc) is a debilitating autoimmune disease featuring immune activation, vasculopathy and auto-antibody production. These processes eventually lead to fibrosis of skin and internal organs (1). The exact aetiology still needs to be unravelled, although it is generally accepted that multiple common genetic variants contribute to the risk of developing SSc (2). Traditionally, SSc has been regarded as a disease propelled by a T helper 2 (Th2) response (3). This view was mainly based on the increased expression of several signature molecules that are associated with Th2 response, such as Interleukin 4 (IL4), Interleukin 13 (IL13) and Interleukin 5 (IL5) in SSc serum, skin and bronchoalveolar lavage (4,5). However, evidence is accumulating that other T-cell subsets including Th1, Th17 and T regulatory cells may also drive pathology observed in SSc (6,7). In parallel with the important role of IL4 and IL13 in Th2 responses, both cytokines seem to exhibit a key role in tuning Th17 responses. For instance, both IL13 and IL4 are able to attenuate Th17 cytokine production (8,9). Adding to the complexity, cells of the innate immune system may also produce pathologically relevant quantities of IL13, especially in the context of fibrosing diseases (10, 11). In addition to being involved in the same biologic processes, IL4 and IL13 display similar features in structure and signalling. Although they share only 25% homology at the amino acid level, their core structure is very similar. The genes are closely situated to each other on chromosome 5q31.1 and are often co-regulated. Both cytokines mediate their effects by interacting with the same receptor complex composed of two transmembrane proteins, IL4RA1 and IL13RA1 (12,13). A second IL13 receptor with a short cytoplasmic tail, IL13RA2, binds IL13 with high affinity and acts as a decoy receptor though a recent study suggests that it may mediate TGFβ induced fibrosis (14).

Polymorphisms of both *IL4* and *IL13* and their receptors have been implicated in susceptibility to asthma (15) and atopic dermatitis and are believed to play a role in psoriasis as well (16,17). In addition, *IL13* variants have recently been implicated in susceptibility to psoriatic arthritis (18). To further elucidate the role of polymorphisms in the *IL4*, *IL13*, *IL4R* and *IL13RA1* genes, we aimed to validate previous associations with immune mediated diseases, including SSc, in a large

multi-national SSc cohort (19-21). Lastly, we investigated the possible effect of these polymorphisms on *IL13*, *IL13RA1* and *IL4RA* expression in B cells, T cells, myeloid dendritic cells (MyDC), plasmacytoid dendritic cells (PDC) and monocytes from SSc patients.

Methods

Patients and Controls

The study population was composed of 2488 SSc patients and 2246 healthy controls matched by geographical region and age. Six case-control sets were of European ancestry; a Spanish cohort: 231 SSc patients and 250 controls; a Dutch cohort: 143 SSc patients and 274 controls; a German cohort: 422 SSc patients and 266 controls; a British cohort: 234 SSc patients and 98 controls: an Italian cohort 444 SSc patients and 362 controls; and a French cohort: 1014 SSc patients and 996 controls. All the patients fulfilled the 1980 American College of Rheumatology (ACR) classification criteria for SSc (22). The local ethical committee from each center approved the study. Both patients and controls were included in the study after written informed consent. All patients included in this study were classified as having limited cutaneous (lcSSc) or diffuse cutaneous SSc (dcSSc) using the criteria postulated by LeRoy (23). Patients with scleroderma changes limited to the skin distal to elbows and/or knees, regardless of facial involvement, fulfill the definition for lcSSc. Those SSc patients with more proximal scleroderma skin changes were classified as having dcSSc. The presence of pulmonary fibrosis was investigated by a high resolution computed tomography scan. Restrictive syndrome and diffusion capacity of the lungs was defined as a forced vital capacity (FVC) < 70% of the predicted value and a diffusion capacity of the lung for carbon monoxide (DLCO) of less than 70% of predicted. Pulmonary artery hypertension was diagnosed by right heart catheterization and considered confirmed if the mean pulmonary artery pressure was greater than 25 mmHg at rest with a normal left atrial wedge pressure. Furthermore, we used follow up data on FVC decline and PAH development from an inception cohort including 358 Dutch and Italian patients, starting inclusion after the onset of the first non-Raynauds symptom and ending at 15 years of follow-up. The patients were evaluated at least yearly for these complications (**Table 1**).

Table 1. Basic and clinical characteristics of the 6 SSc cohorts included in this study.

Population	The Netherlands	Spain	Germany	France	Italy	United Kingdom
Number	143	231	422	1014	444	234
Age (years, (SD))	58 (13)	58 (13)	57(12)	56 (13)	55 (13)	54 (12)
Disease duration (months, (SD))	131 (82)	144 (90)	113 (109)	128 (98)	140 (138)	155 (92)
Female %	81	82	76	86	92	85
Limited phenotype %	69	68	51	63	52	79
Positivity anti-topo %	23	23	26	24	33	15
Positivity ACA %	58	39	46	37	32	71
Pulmonary fibrosis CT scan	32.3	30.7	37.2	39.0	32.1	43.2
Low FVC (<70% predicted) %	26.1	29.1	18.5	16.2	15.3	30.1
Low DLCO (<70% predicted) %	33	45.1	50.2	N.A.	67.2	11.5

Genotyping

Peripheral blood samples (10 ml) collected in EDTA tubes were obtained from each patient and stored at -80°C prior to DNA isolation. Genomic DNA was extracted from leucocytes in peripheral venous blood according to standard protocols. DNA was transferred to 96 wells plates. Each plate contained 3 negative controls (H_2O) and 5 duplicate samples (3 within the plate and 2 between the plates). We chose SNPs that have previously been shown to influence susceptibility to immune mediated diseases and are protein altering, or associated with altered expression levels (19-21). In addition, the *IL13RA1* SNP (rs6646259) covers the largest part of the *IL13RA1* gene since it is situated in a large haplotype block. Taqman assays were used for genotyping SNPs within *IL4* (*−590C>T*/rs2243250/assay ID:C__16176216_10) and the *IL4R* (*Q576R*/rs1801275/assay ID:C___2351160_20). In the *IL13* gene two SNPs were genotyped *R130Q* (rs20541/assay ID:C___2259921_20) and *-1112C>T* (rs1800925/assay ID:C___8932056_10). In the *IL13RA1* gene, the *43163G>A* (rs6646259/assay ID:C__11770516_10) variant was genotyped. Taqman assays were performed according to the protocol from the manufacturer using the 7500 Fast Real-Time PCR system (Applied Biosystems, Foster City, California, USA). Results were analyzed using Sequence Detection Software version 1.4. A competitive allele

specific PCR system (Kaspar Genotyping, Kbioscience, Hoddeston, UK) was used to genotype these SNPs in the French sample as previously reported (24).

Isolation of cell subsets and expression analysis

From 25 SSc patients and 9 controls, derived from the Boston University Medical Center (Boston, MA, USA) we isolated PBMCs from heparinized venous blood by density-gradient centrifugation. Subsequently, BDCA4+ (PDC), CD3+ (T cell), CD19+ (B cell), CD1C+ (MyDC) and CD14+ (monocyte) cells were isolated by magnetic cell separation techniques according to the manufacturer's protocol and as described previously (25). RNA was purified by AllPrep DNA/RNA columns (Qiagen, Valencia, CA), and cDNA was synthesized by I-script (Bio-Rad, Hercules, CA). Quantitative real-time PCR (qRT-PCR) was performed on an Mx3005P QPCR System (Stratagene). Each primer set yielded a product with a dissociation curve composed of a single peak. Ct values for duplicate samples were averaged and the amount of cDNA relative to a housekeeping gene (*GAPDH)* was calculated with the ΔCt method. Primers are displayed in **table 2** and have been derived from the Harvard primerbank or were created with Primer3 software (26,27).

Table 2. Primer sequences used for RT-PCR

Gene	Forward Primers (5'-3')	Reverse Primers (5'-3')	Primerbank ID
GAPDH	ATGGGGAAGGTGAAGGTCG	GGGGTCATTGATGGCAACAATA	7669492a1
IL13	GAAGGCTCCGCTCTGCAAT	TCTGGGTCTTCTCGATGGCA	26787978a1
IL13RA1	ACTCCTGCTTTACCTAAAAAGGC	GCACTACAGAGTCGGTTTCCT	4504647a1
IL4RA	TCATGGATGACGTGGTCAGT	CAGGTCAGCAGCAGAGTGTC	*

*Primer was newly designed with Primer3 software; Rozen S, Skaletsky H.Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol. 2000;132:365-86.

Data analysis

Significance levels were calculated with 2x2 contingency tables and Fisher's exact test by SPSS 16.0. To account for multiple testing the Bonferroni adjustment was applied (significance threshold p =0.002). Homogeneity of OR among cohorts was calculated using Breslow-Day and Woolf Q methods and the calculation of the pooled OR was performed under a fixed-effects model (Mantel-Haenszel meta-analysis). Power calculations using the pooled sample size (2488 SSc patients and

2246 controls) showed that we reached a power of detecting a relative risk of 1.2 of respectively 97% (both rs1801275 and rs20541), 95% (rs1800925), 92% (rs2243250). Since *IL13RA1* is located on the X chromosome we performed an analysis in the female population only (SSc n=2036, controls n=1035), doing this we still reached a power of 84% to detect a relative risk of 1.2 (rs6646259). Survival analysis was performed using Kaplan-Meier curves and significance levels were calculated with Log Rank (Mantel-Cox) statistics. Cox Proportional Hazards Survival Regression was used to determine relative risks. Capped lines represent standard error of the mean throughout the manuscript.

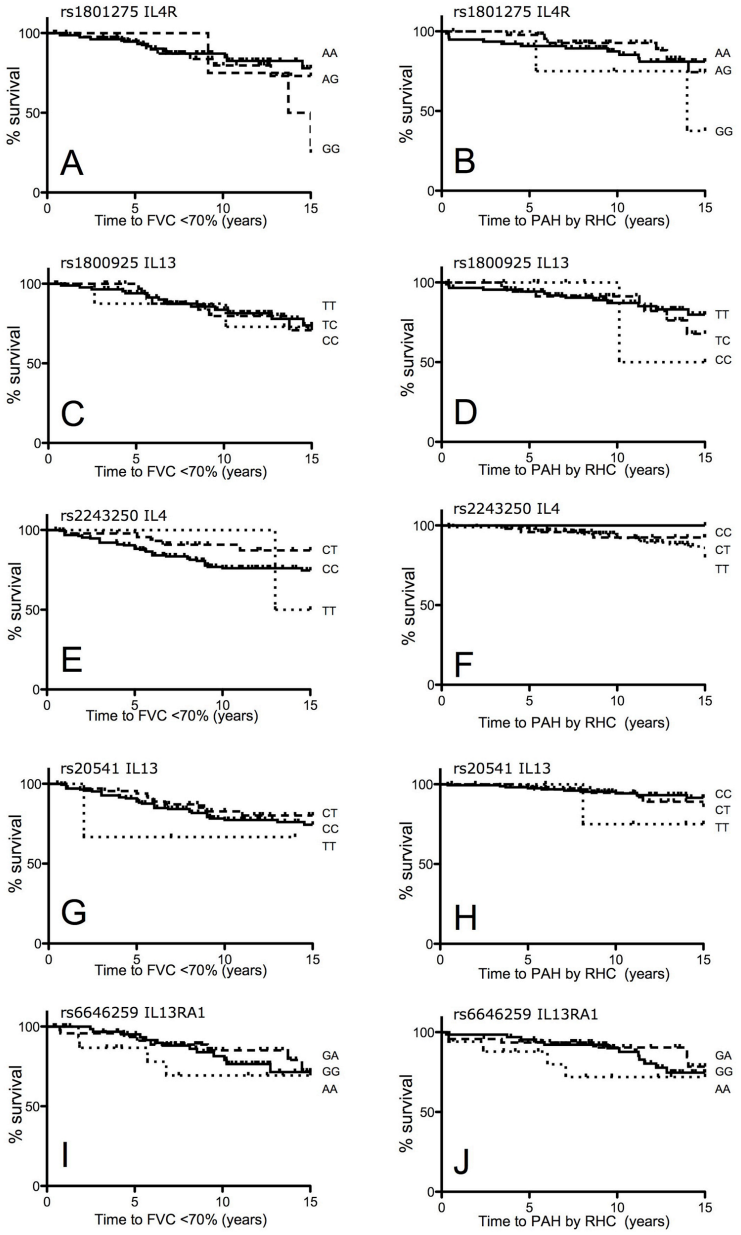
Results

After genotyping, no divergence in Hardy Weinberg Equilibrium was observed. The minor allele frequencies (MAF) of the SNPs tested were all in line with the frequencies reported by the HAPMAP project (www.hapmap.org). The 6 SSc populations showed very little variation in the MAF. According to the Breslow-Day statistics, no significant heterogeneity affecting the six European populations was detected, justifying a meta-analysis. We observed no significant deviation in genotype and allele frequencies in any of the polymorphisms tested in SSc patients compared to controls. Initially, we observed a divergence in allele distribution of the rs1800925 polymorphism in the French dcSSc (p =0.02) and anti-topoisomerase positive (p =0.03) population with SSc compared to controls (**Supplementary Tables 1-5**). However, after correction for multiple testing using the Bonferroni adjustment, no results remained significant. In addition, a meta-analysis taking into account all six European SSc populations did not reveal a significant effect for any of the investigated variants on SSc susceptibility or clinical phenotypes (**Table 3**). When we corrected for gender in our analysis for the non X chromosome located SNPs, no significant differences were observed. Finally, we performed a recessive and dominant analysis which yielded similar negative results (data not shown).

Table 3. Combined analysis of the five investigated polymorphisms, using Mantel-Haenszel analyisis under a fixed model for estimation of combined effects.

SNP	Subtype	total n	minor allele frequency	M-H *p*
IL13 rs1800925	SSc	1832	0.19	0.15
	lcSSc	1125	0.18	0.11
	dcSSc	558	0.20	0.06
	ACA	689	0.18	0.67
	Anti-topo	426	0.20	0.07
	Controls	1869	0.17	
IL13 rs20541	SSc	2474	0.37	0.34
	lcSSc	1520	0.35	0.14
	dcSSc	723	0.41	0.97
	ACA	1046	0.29	0.77
	Anti-topo	563	0.41	0.75
	Controls	2246	0.44	
IL13RA1 rs6646259	SSc	2037	0.19	0.2
(Females only)	lcSSc	1211	0.16	0.78
	dcSSc	556	0.23	0.19
	ACA	800	0.13	0.25
	Anti-topo	413	0.22	0.85
	Controls	1035	0.16	
IL4 rs2243250	SSc	1829	0.16	0.46
	lcSSc	1128	0.16	0.7
	dcSSc	553	0.15	0.3
	ACA	692	0.16	0.81
	Anti-topo	422	0.14	0.26
	Controls	1869	0.15	
IL4R rs1801275	SSc	1819	0.20	0.11
	lcSSc	1125	0.19	0.64
	dcSSc	547	0.20	0.16
	ACA	685	0.19	0.4
	Anti-topo	419	0.20	0.19
	Controls	1879	0.19	

The implication of IL4 and IL13 in many immune mediated pulmonary diseases: (11, 14, 28-32) led us to further investigate the role of functional variants in the IL4 and IL13 genes with special emphasis on pulmonary involvement. For this purpose we used follow up data on FVC decline and PAH development from 358 Dutch and Italian patients, starting at the date of onset of the first non-Raynaud's symptom and ending at 15 years. Patients were evaluated at least once a year for these complications. None of the polymorphisms influenced development of these complications significantly in the follow-up period (**Supplementary Figure 1**).



Supplementary Figure 1. Effect of investigated polymorphisms on pulmonary complication development. No significant effects of the polymorphisms on the development of either a deterioration of forced vital capacity (FVC) below 70% of predicted (A,C,E, G and I) or on the development of pulmonary arterial hypertension (PAH) (B,D,F,H and J) were observed. Both log-rank and hazard ratio's were calculated in 358 patients for a period of 15 years starting at the onset of the first non-Raynaud's symptom (A-J).

To investigate whether polymorphisms in *IL13*, *IL4RA* and *IL13RA1* affect gene expression, the levels of transcripts for these genes were determined by qRT-PCR in purified populations of monocytes, plasmacytoid dendritic cells, myeloid dendritic cells and T cells from the peripheral blood of healthy controls and SSc patients. Overall expression levels of *IL13* were low with either undetectable or very high Ct values in most samples, precluding a comparative analysis for this gene. No statistically significant differences in the expression of *IL13RA1* or *IL4RA* were found in the cell subsets between SSc patients and healthy controls (**Figure 1**). In addition, neither the polymorphism in *IL13RA1* nor the one in *IL4RA* influenced expression of their corresponding gene in any cell type tested (**Figure 2**).

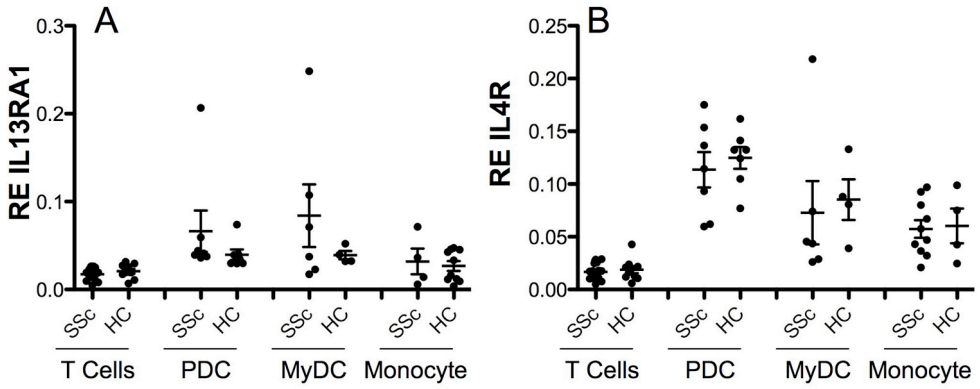


Figure 1. mRNA expression levels of (A) IL13RA1 and (B) IL4RA in T cells, plasmacytoid dendritic cells (PDCs), myeloid dendritic cells (MyDC) and monocytes from healthy controls and SSc patients. No significant differences were observed between healthy controls and SSc patients for either transcript in any cell type. (RE: relative expression compared to GAPDH)

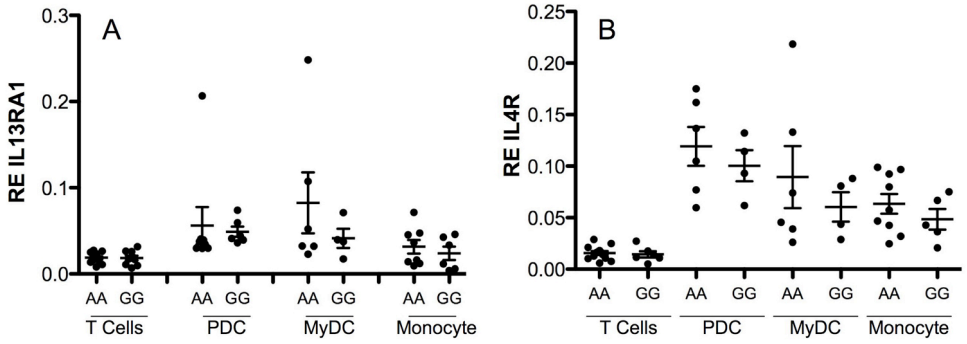


Figure 2. mRNA expression levels of (A) IL13RA1 and (B) IL4RA in T cells, plasmacytoid dendritic cells (PDCs), myeloid dendritic cells (MyDC) segregated by genotype for rs6646259 and rs1801275, respectively. Neither polymorphism significantly influenced gene expression in both SSc patients and controls.

Discussion

In this study we show that five common polymorphisms in the coding regions of *IL4*, *IL13* or their corresponding receptors are not associated with SSc susceptibility in six European populations. The role for both IL4 and IL13 has been addressed in multiple autoimmune diseases including SSc (5, 10, 14-18, 29-36). The functional genetic variants included in this study have been previously associated with immune mediated diseases however they do not seem to contribute to SSc pathogenesis. Furthermore, genetic variation at the *IL13RA1* and *IL4RA* loci did not influence gene expression across a variety of immune cells. However, differences in expression at the protein level or within immune cells of lesional tissue cannot be excluded. On the other hand, a recent genome wide association study which covered these genes fully, did not reveal a statistical significant association either (37). To fully exclude a role for these genes in the pathogenesis of SSc, resequencing would be necessary to cover rare variants as well. It has to be noted that this study did not address a variant in the *IL13RA2* gene previously associated with SSc in a small French cohort (n=107) (38). This variant was left out because a possible overlap between the French cohort in our study and the cohort in the previous report could not be excluded. Considering the size and power of our study to detect significant deviations of all the five variants between cases and controls, the lack of association is improbable due to a type 2 error. The patients included in this study are well-described and evaluated following conventional guidelines (22,23), making disease classification bias unlikely. The patients used in the follow-up cohort were included upon the first evaluation of a non-Raynaud's SSc symptom, flowing from this it may be that patients with extremely progressive disease and with high mortality are not included in this analysis. On the other hand, the time these patients were evaluated covers 15 years and therefore reflects the vast majority of patients that develop pulmonary complications later in disease course, making these results relevant. In SSc mouse models, fibrosis after Bleomycin administration is dependent on IL13 (10). In addition to this, both cytokines are able to drive collagen production by fibroblasts directly (39). However T-cells do not seem to be the key producers of these cytokines and other celltypes are likely to be the main producers of these cytokines in SSc (40). For instance, alveolar macropages have been found to produce this

profibrotic cytokine in pulmonary fibrosis (32). In this study we do not observe statistical significant differences in immune cell subtypes of SSc and HC. However, it has been recently described that scleroderma patients display heterogeneity at the gene expression level that may not be reflected in clinical phenotype (41). We cannot exclude the possibility that a RNA expression level analysis similar to ours across molecular sub-types of SSc might reveal an association especially in patients that show an immune cell activation signature. In addition, genetic variation at the *IL13RA1* and *IL4RA* loci did not influence gene expression across a variety of immune cells. Differences in expression at the protein level or within immune cells of lesional tissue cannot be excluded. Lastly, expression levels of resting cells were measured, whereas differences may become more apparent when placed in a pro-inflammatory environment.

Our results make it likely that the higher levels of IL4 and IL13 observed in SSc patients are not caused by the common genetic variations investigated in this study but may result from other polymorphisms or upstream immune activation mechanisms that promote Th2 maturation or the differentiation of innate immune cells that can also make these cytokines (3,4). For instance, recently aberrances in expression of genes upstream of IL4 and IL13 have been identified in SSc that may have an impact on expression of these genes, instead of polymorphisms in *IL4* and *IL13* (42). In conclusion, we could not replicate the previous association of the rs1800925 polymorphism in the *IL13* gene with SSc, which is most likely due to the small population of SSc patients involved in the initial study (n=107) (43). In addition, we did not observe an association with functional genetic variants of *IL4*, *IL4RA* and *IL13RA1* with SSc susceptibility and/or phenotype. Further studies are necessary to investigate the mechanisms involved in the up regulation of, and response to, IL4 and IL13 observed in SSc.

References

1. Gabrielli A, Avvedimento EV, Krieg T. Scleroderma. N Engl J Med 2009; 360(19):1989-2003.

2. Agarwal SK, Reveille JD. The genetics of scleroderma (systemic sclerosis). Curr Opin Rheumatol 2010; 22(2): 133-8.

3. Mavalia C, Scaletti C, Romagnani P, Carossino AM, Pignone A, Emmi L, Pupilli C, Pizzolo G, Maggi E, Romagnani S. Type 2 helper T-cell predominance and high CD30 expression in systemic sclerosis. Am J Pathol. 1997 Dec;151(6):1751-8.

4. Hancock A, Armstrong L, Gama R, Millar A. Production of interleukin 13 by alveolar macrophages from normal and fibrotic lung. Am J Respir Cell Mol Biol. 1998 Jan;18(1):60-5.

5. Hasegawa M, Fujimoto M, Kikuchi K, Takehara K. Elevated serum levels of interleukin 4 (IL4), IL-10, and IL13 in patients with systemic sclerosis. J Rheumatol. 1997 Feb;24(2):328-32.

6. Radstake TR, van Bon L, Broen J, Wenink M, Santegoets K, Deng Y, Hussaini A, Simms R, Cruikshank WW, Lafyatis R. Increased frequency and compromised function of T regulatory cells in systemic sclerosis (SSc) is related to a diminished CD69 and TGFbeta expression. PLoS One. 2009 Jun 22;4(6):e5981.

7. Radstake TR, van Bon L, Broen J, Hussiani A, Hesselstrand R, Wuttge DM, Deng Y, Simms R, Lubberts E, Lafyatis R The pronounced Th17 profile in systemic sclerosis (SSc) together with intracellular expression of TGFbeta and IFNgamma distinguishes SSc phenotypes. PLoS One. 2009 Jun 17;4(6):e5903.

8. Newcomb DC, Zhou W, Moore ML, Goleniewska K, Hershey GK, Kolls JK, Peebles RS Jr. A functional IL13 receptor is expressed on polarized murine CD4+ Th17 cells and IL13 signaling attenuates Th17 cytokine production. J Immunol. 2009 May 1;182(9):5317-21.

9. Lexberg MH, Taubner A, Förster A, Albrecht I, Richter A, Kamradt T, Radbruch A, Chang HD. Th memory for interleukin-17 expression is stable in vivo. Eur J Immunol. 2008 Oct; 38(10):2654-64.

10. Aliprantis AO, Wang J, Fathman JW, Lemaire R, Dorfman DM, Lafyatis R, Glimcher LH.Transcription factor T-bet regulates skin sclerosis through its function in innate immunity and via IL-13.Proc Natl Acad Sci U S A. 2007 Feb 20;104(8):2827-30. Epub 2007 Feb 16.

11. Kim EY, Battaile JT, Patel AC, You Y, Agapov E, Grayson MH, Benoit LA, Byers DE, Alevy Y, Tucker J, Swanson S, Tidwell R, Tyner JW, Morton JD, Castro M, Polineni D, Patterson GA, Schwendener RA, Allard JD, Peltz G, Holtzman MJ.Persistent activation of an innate immune response translates respiratory viral infection into chronic lung disease.Nat Med. 2008 Jun;14(6): 633-40. Epub 2008 May 18.

12. Hershey GK.IL-13 receptors and signaling pathways: an evolving web.J Allergy Clin Immunol. 2003 Apr;111(4): 677-90; quiz 691.

13. Mueller TD, Zhang JL, Sebald W, Duschl A.Structure, binding, and antagonists in the IL4/IL13 receptor system. Biochim Biophys Acta. 2002 Nov 11;1592(3):237-50.

14. Fichtner-Feigl S, Strober W, Kawakami K, Puri RK, Kitani A.IL-13 signaling through the IL-13alpha2 receptor is involved in induction of TGF-beta1 production and fibrosis.Nat Med. 2006 Jan;12(1):99-106. Epub 2005 Dec 4.

15. Hall IP. Interleukin-4 receptor alpha gene variants and allergic disease. Respir Res. 2000;1(1):6-8. Epub 2000 Jun 23.

16. Elder JT. Genome-wide association scan yields new insights into the immunopathogenesis of psoriasis. Genes Immun. 2009 Apr;10(3):201-9. Epub 2009 Mar

17. Kiyohara C, Tanaka K, Miyake Y. Genetic susceptibility to atopic dermatitis. Allergol Int. 2008 Mar;57(1): 39-56. Epub 2008 Mar 1.

18. Bowes J, Eyre S, Flynn E, Ho P, Salah S, Warren RB, Marzo-Ortega H, Coates L, McManus R, Ryan AW, Kane D, Korendowych E, McHugh N, Fitzgerald O, Packham J, Morgan AW, Griffiths CE, Bruce IN, Worthington J, Barton A. Evidence to support IL-13 as a risk locus for psoriatic arthritis but not psoriasis vulgaris. Ann Rheum Dis. 2011 Jun;70(6):1016-9. Epub 2011 Feb 23.

19. Akkad DA, Arning L, Ibrahim SM, Epplen JT. Sex specifically associated promoter polymorphism in multiple sclerosis affects interleukin 4 expression levels. Genes Immun. 2007 Dec;8(8):703-6. Epub 2007 Sep 13

20. Burgos PI, Causey ZL, Tamhane A, Kelley JM, Brown EE, Hughes LB, Danila MI, van Everdingen A, Conn DL, Jonas BL, Callahan LF, Smith EA, Brasington RD Jr, Moreland LW, van der Heijde DM, Alarcón GS, Bridges SL Jr. Association of IL4R single-nucleotide polymorphisms with rheumatoid nodules in African Americans with rheumatoid arthritis. Arthritis Res Ther. 2010;12(3):R75. Epub 2010 May 5.

21. Miyake Y, Kiyohara C, Koyanagi M, Fujimoto T, Shirasawa S, Tanaka K, Sasaki S, Hirota Y. Department of Public Health, Faculty of Medicine, Fukuoka University, Fukuoka, Japan. Case-Control Study of Eczema Associated with IL13 Genetic Polymorphisms in Japanese Children. Int Arch Allergy Immunol. 2010 Oct 25;154(4):328-335. [Epub ahead of print]

22. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. Arthritis Rheum 1980; 23(5):581-90.

23. LeRoy EC, Black C, Fleischmajer R, Jablonska S, Krieg T, Medsger TA, Jr. et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. J Rheumatol 1988; 15(2):202-5.

24. Dieudé P, Guedj M, Wipff J, Ruiz B, Riemekasten G, Matucci-Cerinic M, Melchers I, Hachulla E, Airo P, Diot E, Hunzelmann N, Cabane J, Mouthon L, Cracowski JL, Riccieri V, Distler J, Meyer O, Kahan A, Boileau C, Allanore Y.Association of the TNFAIP3 rs5029939 variant with systemic sclerosis in the European Caucasian population. Ann Rheum Dis. 2010 Nov; 69(11):1958-64. Epub 2010 May 28.

25. Broen JC, Wolvers-Tettero IL, Geurts-van Bon L, Vonk MC, Coenen MJ, Lafyatis R, Radstake TR, Langerak AW. Skewed X chromosomal inactivation impacts T regulatory cell function in systemic sclerosis. Ann Rheum Dis. 2010 Dec;69(12):2213-6. Epub 2010 Aug 10.

26. Spandidos A, Wang X, Wang H and Seed B. PrimerBank: a resource of human and mouse PCR primer pairs for gene expression detection and quantification. Nucl. Acids Res. 2010 38:D792-9.

27. Rozen S, Skaletsky H.Primer3 on the WWW for general users and for biologist programmers.Methods Mol Biol. 2000;132:365-86.

28. Finotto S, Hausding M, Doganci A, Maxeiner JH, Lehr HA, Luft C, Galle PR, Glimcher LH.Asthmatic changes in mice lacking T-bet are mediated by IL-13.Int Immunol. 2005 Aug;17(8): 993-1007. Epub 2005 Jul 6.

29. Fulkerson PC, Fischetti CA, Hassman LM, Nikolaidis NM, Rothenberg ME.Persistent effects induced by IL-13 in the lung.Am J Respir Cell Mol Biol. 2006 Sep;35(3):337-46. Epub 2006 Apr 27.

30. Keane MP, Gomperts BN, Weigt S, Xue YY, Burdick MD, Nakamura H, Zisman DA, Ardehali A, Saggar R, Lynch JP 3rd, Hogaboam C, Kunkel SL, Lukacs NW, Ross DJ, Grusby MJ, Strieter RM, Belperio JA.IL-13 is pivotal in the fibro-obliterative process of bronchiolitis obliterans syndrome.J Immunol. 2007 Jan 1;178(1):511-9.

31. Yang G, Volk A, Petley T, Emmell E, Giles-Komar J, Shang X, Li J, Das AM, Shealy D, Griswold DE, Li L.Anti-IL-13 monoclonal antibody inhibits airway hyperresponsiveness, inflammation and airway remodeling.Cytokine. 2004 Dec 21;28(6):224-32.

32. Belperio JA, Dy M, Burdick MD, Xue YY, Li K, Elias JA, Keane MP.Interaction of IL-13 and C10 in the pathogenesis of bleomycin-induced pulmonary fibrosis. Am J Respir Cell Mol Biol. 2002 Oct;27(4):419-27.

33. Fuschiotti P, Medsger TA Jr, Morel PA.Effector CD8+ T cells in systemic sclerosis patients produce abnormally high levels of interleukin-13 associated with increased skin fibrosis.Arthritis Rheum. 2009 Apr;60(4):1119-28.

34. Hasegawa M, Sato S, Nagaoka T, Fujimoto M, Takehara K.Serum levels of tumor necrosis factor and interleukin-13 are elevated in patients with localized scleroderma.Dermatology. 2003;207(2):141-7.

35. Riccieri V, Rinaldi T, Spadaro A, Scrivo R, Ceccarelli F, Franco MD, Taccari E, Valesini G.Interleukin-13 in systemic sclerosis: relationship to nailfold capillaroscopy abnormalities.Clin Rheumatol. 2003 May;22(2):102-6.

36. Liu T, Jin H, Ullenbruch M, Hu B, Hashimoto N, Moore B, McKenzie A, Lukacs NW, Phan SH. Regulation of found in inflammatory zone 1 expression in bleomycin-induced lung fibrosis: role of IL-4/IL-13 and mediation via STAT-6. J Immunol. 2004 Sep 1;173(5):3425-31.

37. Radstake TR, Gorlova O, Rueda B, Martin JE, Alizadeh BZ, Palomino-Morales R, Coenen MJ, Vonk MC, Voskuyl AE, Schuerwegh AJ, Broen JC, van Riel PL, van 't Slot R, Italiaander A, Ophoff RA, Riemekasten G, Hunzelmann N, Simeon CP, Ortego-Centeno N, González-Gay MA, González-Escribano MF; Spanish Scleroderma Group, Airo P, van Laar J, Herrick A, Worthington J, Hesselstrand R, Smith V, de Keyser F, Houssiau F, Chee MM, Madhok R, Shiels P, Westhovens R, Kreuter A, Kiener H, de Baere E, Witte T, Padykov L, Klareskog L, Beretta L, Scorza R, Lie BA, Hoffmann-Vold AM, Carreira P, Varga J, Hinchcliff M, Gregersen PK, Lee AT, Ying J, Han Y, Weng SF, Amos CI, Wigley FM, Hummers L, Nelson JL, Agarwal SK, Assassi S, Gourh P, Tan FK, Koeleman BP, Arnett FC, Martin J, Mayes MD. Genome-wide association study of systemic sclerosis identifies CD247 as a new susceptibility locus.Nat Genet. 2010 May;42(5): 426-9. Epub 2010 Apr 11.

38. Granel B, Allanore Y, Chevillard C, Arnaud V, Marquet S, Weiller PJ, Durand JM, Harlé JR, Grange C, Frances Y, Berbis P, Gaudart J, de Micco P, Kahan A, Dessein A. IL13RA2 gene polymorphisms are associated with systemic sclerosis. J Rheumatol. 2006 Oct;33(10):2015-9. Epub 2006 Sep 15.

39. Bhogal RK, Bona CA. Regulatory effect of extracellular signal-regulated kinases (ERK) on type I collagen synthesis in human dermal fibroblasts stimulated by IL-4 and IL-13. Int Rev Immunol. 2008;27(6):472-96.

40. Helene M, Lake-Bullock V, Zhu J, Hao H, Cohen DA, Kaplan AM. T cell independence of bleomycin-induced

pulmonary fibrosis. J Leukoc Biol. 1999 Feb;65(2):187-95.

41. Milano A, Pendergrass SA, Sargent JL, George LK, McCalmont TH, Connolly MK, Whitfield ML.Molecular subsets in the gene expression signatures of scleroderma skin. PLoS One. 2008 Jul 16;3(7):e2696.

42. Katsumoto TR, Whitfield ML, Connolly MK.The pathogenesis of systemic sclerosis. Annu Rev Pathol. 2011 Feb 28;6:509-37.

43. Granel B, Chevillard C, Allanore Y, Arnaud V, Cabantous S, Marquet S, Weiller PJ, Durand JM, Harlé JR, Grange C, Frances Y, Berbis P, Gaudart J, de Micco P, Kahan A, Dessein A. Evaluation of interleukin 13 polymorphisms in systemic sclerosis.Immunogenetics. 2006 Aug; 58(8):693-9. Epub 2006 Jul 11.

Supplementary Tables.

Supplementary Table 1: Overview of the genotype and allele frequencies of *IL13* rs1800925 in the individual populations

IL13 rs1800925	Phenotype	n	TT	TC	CC	*p*	MAF	*p*
Netherlands	**SSc**	143	0.69	0.27	0.03	0.21	0.17	0.78
	lcSSc	98	0.67	0.30	0.03	0.51	0.18	0.99
	dcSSc	33	0.68	0.30	0.02	0.76	0.17	0.87
	ACA	84	0.68	0.29	0.04	0.38	0.18	0.99
	Anti-topo	33	0.66	0.30	0.04	0.76	0.19	0.74
	Controls	274	0.65	0.33	0.01		0.18	
Spain	**SSc**	231	0.62	0.33	0.04	0.015	0.21	0.12
	lcSSc	157	0.64	0.34	0.02	0.31	0.19	0.45
	dcSSc	56	0.64	0.32	0.04	0.092	0.2	0.49
	ACA	89	0.66	0.33	0.02	0.28	0.18	0.82
	Anti-topo	N.A.	N.A.	N.A.	N.A.		N.A.	
	Controls	250	0.66	0.34	0.00		0.17	
Italy	**SSc**	444	0.59	0.41	0.01	0.08	0.21	0.05
	lcSSc	231	0.64	0.35	0.02	0.15	0.19	0.39
	dcSSc	154	0.67	0.32	0.01	0.36	0.17	0.99
	ACA	144	0.67	0.32	0.01	0.76	0.17	0.99
	Anti-topo	146	0.67	0.33	0.01	0.79	0.17	0.99
	Controls	362	0.66	0.33	0.00		0.17	
France	**SSc**	1014	0.66	0.31	0.03	0.25	0.19	0.22
	lcSSc	639	0.67	0.30	0.03	0.65	0.18	0.58
	dcSSc	315	0.61	0.35	0.03	0.35	0.21	0.02
	ACA	372	0.66	0.32	0.02	0.15	0.18	0.63
	Anti-topo	247	0.61	0.35	0.04	0.06	0.21	0.03
	Controls	983	0.69	0.28	0.03		0.17	

Supplementary Table 2: Overview of the genotype and allele frequencies of *IL13* rs20541 in the individual populations

IL13 rs20541	Phenotype	n	CC	CT	TT	p	MAF	p
Netherlands	**SSc**	143	0.64	0.30	0.06	0.91	0.21	0.79
	lcSSc	98	0.69	0.29	0.03	0.61	0.17	0.34
	dcSSc	33	0.64	0.27	0.08	0.7	0.22	0.63
	ACA	84	0.64	0.32	0.04	0.71	0.20	0.99
	Anti-topo	33	0.64	0.27	0.08	0.7	0.22	0.63
	Controls	274	0.65	0.29	0.05		0.20	
Spain	**SSc**	231	0.64	0.29	0.06	0.95	0.21	0.75
	lcSSc	157	0.68	0.29	0.04	0.38	0.18	0.21
	dcSSc	56	0.64	0.29	0.08	0.96	0.22	0.99
	ACA	89	0.63	0.29	0.07	0.93	0.22	0.99
	Anti-topo	N.A.	N.A.	N.A.	N.A.		N.A.	
	Controls	250	0.63	0.30	0.07		0.22	
Italy	**SSc**	444	0.66	0.30	0.04	0.72	0.19	0.61
	lcSSc	231	0.66	0.30	0.04	0.75	0.19	0.71
	dcSSc	154	0.65	0.30	0.05	0.99	0.20	0.99
	ACA	144	0.67	0.30	0.03	0.48	0.18	0.54
	Anti-topo	146	0.66	0.29	0.06	0.98	0.20	0.99
	Controls	362	0.65	0.30	0.05		0.20	
UK	**SSc**	234	0.64	0.29	0.06	0.7	0.21	0.60
	lcSSc	186	0.67	0.29	0.03	0.92	0.18	0.82
	dcSSc	41	0.62	0.32	0.06	0.68	0.22	0.62
	ACA	167	0.66	0.29	0.04	0.99	0.19	0.99
	Anti-topo	34	0.64	0.29	0.06	0.9	0.21	0.73
	Controls	98	0.66	0.30	0.04		0.19	
Germany	**SSc**	422	0.66	0.29	0.04	0.99	0.19	0.37
	lcSSc	215	0.63	0.30	0.07	0.6	0.22	0.69
	dcSSc	132	0.65	0.30	0.05	0.97	0.20	0.78
	ACA	196	0.64	0.31	0.06	0.82	0.21	0.99
	Anti-topo	109	0.66	0.29	0.04	0.98	0.19	0.55
	Controls	266	0.24	0.11	0.02		0.21	
France	**SSc**	1000	0.66	0.30	0.04	0.082	0.19	0.61
	lcSSc	633	0.67	0.29	0.04	0.079	0.19	0.88
	dcSSc	307	0.64	0.33	0.04	0.412	0.20	0.36
	ACA	366	0.68	0.29	0.03	0.45	0.18	0.70
	Anti-topo	241	0.63	0.34	0.03	0.71	0.20	0.42
	Controls	996	0.65	0.32	0.03		0.19	

Supplementary Table 3: Overview of the genotype and allele frequencies of *IL13RA1* rs6646259 in the individual populations (females only)

IL13RA1 rs6646259		n	GG	GA	AA	*p*	MAF	*p*
Netherlands	**SSc**	116	0.56	0.35	0.08	0.09	0.26	0.15
	lcSSc	79	0.56	0.35	0.08	0.12	0.26	0.22
	dcSSc	27	0.58	0.34	0.08	0.53	0.25	0.59
	ACA	68	0.59	0.35	0.05	0.59	0.23	0.63
	Anti-topo	27	0.57	0.37	0.05	0.95	0.24	0.60
	Controls	222	0.61	0.36	0.03		0.21	
Spain	**SSc**	187	0.57	0.36	0.07	0.54	0.25	0.31
	lcSSc	127	0.58	0.36	0.06	0.75	0.24	0.57
	dcSSc	45	0.58	0.33	0.08	0.49	0.25	0.49
	ACA	72	0.60	0.35	0.06	0.93	0.23	0.82
	Anti-topo		N.A.	N.A.	N.A.		N.A.	
	Controls	203	0.60	0.36	0.04		0.22	
Italy	**SSc**	360	0.58	0.37	0.06	0.68	0.24	0.69
	lcSSc	187	0.60	0.36	0.04	0.9	0.22	0.75
	dcSSc	125	0.56	0.38	0.06	0.78	0.25	0.59
	ACA	117	0.57	0.39	0.03	0.93	0.23	0.99
	Anti-topo	118	0.56	0.41	0.04	0.84	0.24	0.78
	Controls	293	0.58	0.38	0.04		0.23	
UK	**SSc**	190	0.58	0.36	0.06	0.8	0.24	0.66
	lcSSc	151	0.59	0.37	0.05	0.09	0.23	0.91
	dcSSc	33	0.57	0.36	0.07	0.87	0.25	0.61
	ACA	135	0.56	0.40	0.04	0.89	0.24	0.72
	Anti-topo	28	0.61	0.33	0.07	0.73	0.23	0.85
	Controls	79	0.60	0.37	0.04		0.22	
Germany	**SSc**	342	0.60	0.37	0.04	0.89	0.22	0.71
	lcSSc	93	0.59	0.38	0.03	0.97	0.22	0.83
	dcSSc	83	0.56	0.40	0.04	0.97	0.24	0.83
	ACA	63	0.57	0.38	0.05	0.93	0.24	0.90
	Anti-topo	41	0.56	0.42	0.02	0.88	0.23	0.89
	Controls	215	0.58	0.39	0.04		0.23	
France	**SSc**	843	0.52	0.40	0.09	0.75	0.28	0.62
	lcSSc	574	0.53	0.39	0.08	0.77	0.27	0.48
	dcSSc	243	0.50	0.40	0.10	0.69	0.30	0.86
	ACA	345	0.52	0.40	0.08	0.746	0.28	0.63
	Anti-topo	200	0.47	0.44	0.10	0.537	0.31	0.39
	Controls	22	0.59	0.32	0.09		0.25	

Supplementary Table 4: Overview of the genotype and allele frequencies of *IL4* rs2243250 in the individual populations

IL4 rs2243250	Phenotype	n	CC	CT	TT	*p*	MAF	*p*
Netherlands	**SSc**	143	0.74	0.24	0.02	0.67	0.14	0.48
	lcSSc	98	0.73	0.24	0.03	0.96	0.15	0.73
	dcSSc	33	0.75	0.21	0.03	0.89	0.14	0.72
	ACA	84	0.72	0.24	0.04	0.99	0.16	0.99
	Anti-topo	33	0.75	0.21	0.03	0.9	0.14	0.72
	Controls	274	0.72	0.24	0.04		0.16	
Spain	**SSc**	231	0.71	0.25	0.05	0.82	0.17	0.66
	lcSSc	157	0.72	0.24	0.04	0.98	0.16	0.99
	dcSSc	56	0.75	0.21	0.03	0.87	0.14	0.77
	ACA	89	0.72	0.24	0.04	0.91	0.16	0.99
	Anti-topo	N.A.	N.A.	N.A.	N.A.		N.A.	
	Controls	250	0.72	0.25	0.04		0.16	
Italy	**SSc**	444	0.71	0.26	0.03	0.87	0.16	0.99
	lcSSc	231	0.73	0.23	0.05	0.83	0.16	0.99
	dcSSc	154	0.72	0.26	0.02	0.6	0.15	0.71
	ACA	144	0.72	0.24	0.04	0.96	0.16	0.99
	Anti-topo	146	0.73	0.25	0.03	0.89	0.15	0.78
	Controls	362	0.72	0.25	0.04		0.16	
France	**SSc**	1011	0.72	0.25	0.03	0.9	0.15	0.84
	lcSSc	642	0.71	0.26	0.02	0.75	0.15	0.67
	dcSSc	310	0.72	0.26	0.02	0.47	0.15	0.87
	ACA	375	0.71	0.26	0.03	0.86	0.16	0.67
	Anti-topo	243	0.73	0.26	0.01	0.38	0.14	0.59
	Controls	983	0.73	0.24	0.03		0.15	

Supplementary Table 5: Overview of the genotype and allele frequencies of *IL4R* rs1801275 in the individual populations

IL4R rs1801275	Phenotype	n	GG	GA	AA	*p*	MAF	*p*
Netherlands	**SSc**	143	0.67	0.29	0.05	0.96	0.19	0.71
	lcSSc	98	0.66	0.30	0.04	0.91	0.19	0.75
	dcSSc	33	0.66	0.27	0.06	0.94	0.2	0.99
	ACA	84	0.67	0.30	0.03	0.84	0.18	0.58
	Anti-topo	33	0.66	0.27	0.06	0.93	0.2	0.99
	Controls	274	0.65	0.30	0.05		0.2	
Spain	**SSc**	231	0.64	0.29	0.06	0.14	0.21	0.12
	lcSSc	157	0.66	0.29	0.04	0.65	0.19	0.45
	dcSSc	56	0.67	0.27	0.07	0.28	0.2	0.49
	ACA	89	0.69	0.26	0.05	0.43	0.18	0.82
	Anti-topo	N.A.	N.A.	N.A.	N.A.		N.A.	
	Controls	250	0.69	0.29	0.03		0.17	
Italy	**SSc**	444	0.65	0.29	0.07	0.36	0.21	0.38
	lcSSc	231	0.67	0.29	0.03	0.85	0.18	0.65
	dcSSc	154	0.64	0.28	0.08	0.3	0.22	0.27
	ACA	144	0.67	0.28	0.05	0.97	0.19	0.99
	Anti-topo	146	0.65	0.27	0.07	0.36	0.21	0.54
	Controls	362	0.66	0.29	0.04		0.19	
France	**SSc**	1001	0.66	0.29	0.04	0.94	0.19	0.87
	lcSSc	639	0.66	0.30	0.03	0.51	0.19	0.91
	dcSSc	304	0.67	0.27	0.06	0.63	0.19	0.87
	ACA	368	0.65	0.31	0.04	0.77	0.20	0.58
	Anti-topo	240	0.65	0.30	0.05	0.88	0.20	0.61
	Controls	993	0.67	0.29	0.04		0.19	

Chapter 7

A rare polymorphism in *Toll Like Receptor 2* is associated with systemic sclerosis phenotype and increases production of inflammatory mediators.

Broen JCA, Bossini-Castillo L, Van Bon L, Vonk MC, Knaapen H, Beretta L, Rueda B, Hesselstrand R, Herrick A, Worthington J, Hunzelman N, Denton C, Fonseca C, Riemekasten G, Kiener H, Scorza PhD, Simeon CP, Ortego-Centeno N (for the Spanish Systemic Sclerosis group), Gonzalez-Gay MA, Airo' P, Coenen MJH, Martin J and Radstake TRDJ

Arthritis and Rheumatism 2011

Abstract

Aim: To investigate whether polymorphisms in *toll like receptor (TLR)* genes, previously reported to be associated with immune mediated diseases are implicated in systemic sclerosis (SSc).

Methods: We genotyped 14 polymorphisms in the *TLR 2, 4, 7, 8* and *9* genes in a discovery cohort comprising 452 SSc patients and 537 controls and a replication cohort consisting of 1170 SSc patients and 925 controls. Furthermore we analyzed 15 year follow-up data from 964 patients to assess the potential association of *TLR* variants with the development of disease complications. Next to this, we analyzed the functional impact of the associated polymorphism on monocyte derived dendritic cells.

Results: Exploiting the discovery cohort, we observed that a rare functional polymorphism in *TLR2 (Pro631His)*, was associated with anti-topoisomerase positivity (p =0.003 OR 2.24 95%CI:1.24-4.04). This observation was validated in the replication cohort (p=0.0001 OR 2.73 95%CI:1.85-4.04). In addition, the replication cohort also revealed an association between the *TLR2* variant with the diffuse subform of the disease and the development of pulmonary arterial hypertension, respectively (p=0.02, Log-Rank p=0.003, Cox proportional hazards ratio: 5.61 ((95%CI 1.53-20.58)). Functional analysis revealed that monocyte derived dendritic cells carrying the Pro63His variant produce more inflammatory mediators (TNFalpha and IL-6) upon TLR2 mediated stimulation (both p<0.0001).

Conclusion: The rare *TLR2* Pro631His variant is robustly associated with anti-topoisomerase positivity, diffuse SSc and the development of PAH. Besides, this variant influences TLR2 mediated cell responses. Further research is necessary to reveal the precise role of TLR2 in the disease pathogenesis of SSc.

Introduction

Systemic sclerosis (SSc) bares all the hallmarks of an autoimmune disease, which features vasculopathy, immune activation and ultimately extensive fibrosis of skin and internal organs. Although the research in the field of SSc is intensified over the last years, there is still no clear view on the pathogenesis or a cure (1). It is generally accepted that genetic factors play a role in this disease, which is supported by the observation of both familial and ethnic aggregation and the numerous reports describing an association between genetic variants and SSc (2).

The family of pattern recognition receptors (PRPs), and more specifically the family of Toll Like Receptors (TLRs) are among the most scrutinized molecules in immunity and autoimmunity. TLRs provide a first line pathogen recognition system and are able to direct the innate immune system towards the appropriate immune responses (3). Although the main purpose of these receptors is to identify microbial antigens and subsequently mount a proper response, growing evidence points towards their direct implication in various autoimmune diseases (4).

The role for several TLRs has been shown in SSc. For instance, monocyte-derived and myeloid dendritic cells (DCs) from SSc patients display an augmented response to various TLR specifc ligands, some of which have been show to be present in serum of SSc patients (5,6). In addition, subcutaneous administration of TLR ligands in an experimental model provokes a marked inflammation in the skin that partly resembles SSc skin changes (7). However, the extent of the role of the different TLR subsets in SSc has not been investigated intensely. Genetic studies are able to shed light on this latter aspect. Therefore we investigated 14 polymorphisms covering the *TLR 2, 4, 7, 8* and *9* genes, previously associated with immune mediated disease, for their role in SSc susceptibility and corresponding clinical phenotypes. The selection of these SNPs was founded on a Pubmed based search for TLR polymorphisms associated with immune-mediated disease, most preferably with functional impact on the protein or cell function level as well.

Methods

Patients and Controls

We used a discovery cohort composed of 452 SSc patients and 537 healthy controls matched by geographical region, age and gender. This population was composed of three case-control sets of European ancestry; a Spanish cohort: 188 SSc patients and 193 controls; a Dutch cohort: 85 SSc patients and 255 controls; and an Italian cohort: 179 SSc patients and 89 controls. As a replication cohort we exploited a second cohort comprising 1170 SSc patients and 925 controls. This cohort consisted of a Swedish cohort: 193 SSc patients and 167 controls: a German cohort of 312 SSc patients and 247 controls; a second Italian cohort of 158 patients and 231 controls: a second Dutch cohort of 238 SSc patients and 196 controls and an English cohort consisting of 269 SSc patients and 84 controls. Both the allele frequencies of the controls originating from the United Kingdom and Dutch replication cohort were derived from literature reports (8,9).

All patients fulfilled the 1980 American College of Rheumatology (ACR) classification criteria for SSc (10). The local ethical committee from each centre approved the study. Both patients and controls were included in the study after written informed consent. All patients included in this study were classified as having limited cutaneous (lcSSc) or diffuse cutaneous SSc (dcSSc) using the criteria postulated by LeRoy (11). Autoantibody testing was performed using either ELISA or immunofluorescence microscopy. The presence of pulmonary fibrosis was investigated by a high-resolution computed tomography scan. Restrictive syndrome and diffusion capacity of the lungs was defined as a forced vital capacity (FVC) < 70% of the predicted value and a diffusion capacity of the lung for carbon monoxide (DLCO) of less than 70% of predicted. Pulmonary artery hypertension was diagnosed by right heart catheterization and considered confirmed if the mean pulmonary artery pressure was greater than 25 mmHg at rest with a normal left atrial wedge pressure. Furthermore, we used follow up data on FVC and DLCO decline, development of pulmonary fibrosis, kidney involvement and pulmonary arterial hypertension (PAH) development from 964 patients, starting at the date of onset of the first non-Raynauds symptom and ending at 15 years or death. The

patients were evaluated at least yearly for these complications (**supplementary table 1**).

Genotyping

We assessed 14 functional polymorphisms in the *TLR2 ,4, 7, 8* and *9* genes for their role in SSc susceptibility and clinical phenotype. Genotyping was performed using 5' nuclease (Taqman) assays with predesigned primers and probes (Applied Biosystems, Foster City, CA). The polymorphisms investigated and their functional properties and previous associations are displayed in **table 1**.

Table 1. Overview of tested TLR polymorphisms

SNP	Gene	Functional change	AB Assay ID	Previous (immune system related) association	References
rs1898830	TLR2	-15607A>G	C__11853988_10	Bronchiolitis Obliterans	20
rs5743704	TLR2	Pro631His	C__25607736_10	IBD, impaired membrane internalization, acute reactive arthritis	14,21,22
rs4986790	TLR4	896A>G (D299G)	C__11722238_20	Asthma and atopy	23
rs4986791	TLR4	1196C>T p.T399I	C__11722237_20	Liver cirrosis	23
rs7873784	TLR4	16649G>C	C__29292008_10	Multiple sclerosis	25
rs3853839	TLR7	3'UTR	C___2259573_10	SLE, increased expression	26
rs179008	TLR7	Gln11Leu	C___2259574_10	Asthma	27
rs2302267	TLR7		C__15757400_10	Chronic HCV infection	28
rs5743781	TLR7	A448V	C__25643238_10	Newly found variant	29
rs3764879	TLR8	TLR8-129G>C	C___2183829_10	Fatal crimean congo hemorrhagic fever	30
rs3764880	TLR8	Met1Val	C___2183830_10	Isoform regulating, tuberculosis	31,32
rs5741883	TLR8		C__29409072_10	Rheumatoid factor positivity	33
rs5744088	TLR8		C__32184097_10		
rs187084	TLR9		C___2301952_10	Graves' ophthalmopathy	34

Cell isolation, stimulation and expression analysis

Peripheral blood mononuclear cells (PBMCs) were isolated from heparinized venous blood by using density-gradient centrifugation over Ficoll-Paque (Amersham Bioscience). Monocytes were obtained using CD14 and BDCA1 microbeads (Miltenyi Biotec Inc, CA, USA) for monocytes and myeloid DCs, respectively. Generation of monocyte-derived DCs (moDCs) has been extensively described by our group previously (5).

Flowcytometric analysis of CD14, CD86 and MHC expression on moDCs

Phenotypical analysis of moDCs was performed using standardized flow cytometry protocols as described previously (17). DCs were characterized by staining with mAbs against human CD14 (Miltenyi Biotec Inc., Ca, USA), CD86 (BD Bioscience, NJ, USA), MHC-II DR/DP (clone Q1514). Cells were analyzed with a fluorescence-activated cell sorter (FACSCalibur; BD Biosciences) and analyzed with FlowJo 8.7.3, (Treestar, OR, USA) for the proportion of positive cells and the mean fluorescence intensity relative to cells stained with the relevant IgG isotype controls.

Stimulation of moDC with TLR ligands

At day 6 of culture moDCs were re-plated in a concentration of $0.5x10^6$ DCs/ml and either transferred to 24 well (1 ml) or 96 well (100 µl) culture plates. The culture medium consisted of RPMI-1640 Dutch modification (Invitrogen Life Technologies) supplemented with 10% FCS and antibiotic-antimycotic (Invitrogen Life Technologies) in the presence of IL4 (500 U/ml; Schering-Plough) and GM-CSF (800 U/ml; Schering-Plough). Cells were then stimulated with TLR2 agonists for 24 hours, after which supernatants were collected. The concentration of TLR2 ligand Pam3Cys (TLR2, 5 µg/ml, EMC microcollections, Germany) was similar throughout the studies.

Measurement of cytokines

Levels of tumour necrosis factor alpha (TNFa) and Interleukin 6 (IL6) were measured in the supernatants using commercially available kits (Bio-Rad) according to the manufacturer's instructions. Cytokine levels were measured and analyzed with the Bio-Plex system (Bio-Rad). The sensitivity of the cytokine assay was < 5 pg/ml for all cytokines measured.

Statistical analysis.

Statistical analyses were performed using 2x2 contingency tables and Fisher's exact test by using SPSS 16.0. Homogeneity of OR among cohorts was calculated using Breslow-Day and Woolf Q methods and the calculation of the pooled OR was performed under a fixed-effects model (Mantel-Haenszel meta-analysis). Candidate loci themselves were used to test for population stratification, as described previously (12). For stratification tests, when more SNPs were in linkage

disequilibrium (LD), the SNP/LD block with the highest $\chi 2$ test value, that is the SNP that the contributes most to differences across ethnicities, was chosen for the analysis. The *TLR7* and *8* polymorphisms were analyzed in the female population only, since both genes are situated on the X chromosome. Power analysis showed that his study reaches a power of 80% or more to detect the effect of all included polymorphisms at an OR of 1.3 in a log additive model. Survival analysis was performed using Kaplan-Meier curves and significance levels were calculated with Log Rank (Mantel-Cox) statistics. Cox Proportional Hazards Survival Regression was used to determine relative risks. Capped lines represent standard error of the mean throughout the manuscript. For the cellular experiments, values are shown throughout the paper as mean ± sem. Statistical analyses were applied using Students t-test or Mann-Whitney U-test were appropriate. Differences were considered significant for $*p<0.05$, $**p<0.01$ and $***p<0.001$. We did not apply correction for multiple testing in the discovery cohort, to avoid increasing the risk of false negative results not being profoundly investigated in the replication cohort.

Results

Discovery cohort

All the 14 investigated variants were in Hardy-Weinberg equilibrium and frequencies in the control population were very similar to those reported in the HAPMAP project. No significant evidence of population stratification was present. We did not find any significant associations, except for the *TLR2* Pro631His (rs5743704) polymorphism, which was significantly associated with having positive anti-topoisomerase autoantibody titers in both the Dutch and the Spanish discovery cohort (respectively p=0.01 OR: 4.45; 95%CI:1.36-12.54 and p=0.02 OR: 3.01 95%CI:1.07-7.77). When we performed a meta-analysis we found a significant effect for all 3 populations together as well (p=0.003 OR: 2.24 95%CI:1.24-4.04) (**table 2, supplementary Tables 2-14**).

Table 2. Distribution of the genotypes and alleles of the *TLR2* variant (Pro631His) in the discovery cohort

cohort	phenotype	n	CC	CA	AA	*p*	A	*p*
Italy	SSC	165	0.95	0.05	0.00	0.51	0.97	0.6
	LcSSc	116	0.94	0.06	0.00	0.51	0.97	0.79
	DcSSc	40	0.98	0.03	0.00	0.57	0.99	0.44
	ACA	60	0.97	0.03	0.00	0.56	0.98	0.32
	Anti-Topo	74	0.93	0.07	0.00	0.99	0.97	0.99
	Controls	88	0.93	0.06	0.01		0.96	
The Netherlands	SSC	82	0.89	0.11	0.00	0.35	0.95	0.52
	LcSSc	63	0.89	0.11	0.00	0.41	0.94	0.63
	DcSSc	18	0.94	0.06	0.00	0.91	0.97	0.99
	ACA	23	0.96	0.04	0.00	0.82	0.98	0.71
	Anti-Topo	20	0.70	0.30	0.00	**0.01**	0.85	***0.01***
	Controls	249	0.93	0.07	0.00		0.96	
Spain	SSC	186	0.94	0.06	0.00	0.34	0.97	0.35
	LcSSc	129	0.92	0.07	0.01	0.48	0.96	0.52
	DcSSc	51	0.90	0.08	0.02	0.15	0.94	0.24
	ACA	74	0.96	0.04	0.00	0.45	0.98	0.57
	Anti-Topo	31	0.81	0.16	0.03	**0.01**	0.89	***0.02***
	Controls	181	0.92	0.07	0.01		0.95	

cohort	phenotype	n	CC	CA	AA	*p*	A	*p*
Mantel-Haenszel	SSC	433						0.55
Meta-analysis	LcSSc	308						0.68
	DcSSc	109						0.22
	ACA	157						0.21
	Anti-Topo	125						***0.003***
	Controls	518						

Replication cohort.

Since the allele frequency of the associated polymorphism was relatively low (average 3%), we used a large replication set consisting of an additional 1170 SSc patients and 925 controls to validate the findings. This replication cohort was composed of an additional Italian and Dutch population, but also from different ethnic cohorts encompassing German, Swedish and British SSc patients and healthy controls. In this replication cohort we found an association with positive anti-topoisomerase autoantibodies and the *TLR2* Pro631His variant in the Italian, German and Dutch population (respectively *p*=0.02 (OR: 3.15 95%CI:1.04-8.68), *p*=0.001 (OR: 3.72 95%CI:1.64-6.62), *p*=0.02 (OR:3.26 95%CI:1.07-8.93) (**Table 3**). When we combined the results from the replication cohort in a meta-analysis we found a strong overall association (*p*=0.0001 OR: 2.73 95%CI:1.85-4.04). In addition, an association with the dcSSc subtype of the disease was observed, which fits the observation that anti-topoisomerase positivity is usually found in DcSSc (*p*=0.02 OR: 1.67 95%CI:1.08-2.58) (**table 3**).

Table 3. Distribution of the *TLR2* genotypes and alleles of the Pro631His variant in the replication cohort

cohort	phenotype	n	CC	CA	AA	A	p
Italy	SSc	158	0.94	0.06	0.00	0.03	0.99
	LcSSc	97	0.95	0.05	0.00	0.03	0.99
	DcSSc	44	0.89	0.11	0.00	0.06	0.15
	ACA	88	0.95	0.05	0.00	0.03	0.79
	Anti-Topo	39	0.85	0.13	0.03	0.09	***0.02***
	Control	231	0.94	0.06	0.00	0.03	

cohort	phenotype	n	CC	CA	AA	A	p
Sweden	SSc	193	0.92	0.08	0.00	0.04	0.55
	Limited	117	0.94	0.06	0.00	0.03	0.99
	Diffuse	50	0.86	0.14	0.00	0.07	0.08
	ACA	109	0.95	0.05	0.00	0.03	0.79
	Anti-Topo	48	0.88	0.10	0.02	0.07	0.07
	Control	167	0.94	0.06	0.00	0.03	
Germany	SSc	312	0.94	0.06	0.00	0.03	0.74
	Limited	163	0.93	0.07	0.00	0.04	0.99
	Diffuse	121	0.89	0.11	0.00	0.06	0.24
	ACA	146	0.93	0.07	0.00	0.04	0.99
	Anti-Topo	116	0.79	0.19	0.02	0.10	***0.001***
	Control	247	0.93	0.07	0.00	0.04	
The Netherlands	SSc	238	0.97	0.03	0.00	0.02	0.05
	Limited	66	0.95	0.05	0.00	0.03	0.58
	Diffuse	31	0.88	0.12	0.00	0.06	0.31
	ACA	61	0.95	0.05	0.00	0.03	0.58
	Anti-Topo	29	0.86	0.14	0.00	0.11	***0.02***
	Control	196	0.92	0.08	0.00	0.04	
United Kingdom	SSc	269	0.94	0.06	0.00	0.03	0.33
	Limited	172	0.96	0.04	0.00	0.02	0.1
	Diffuse	59	0.89	0.11	0.00	0.06	0.99
	ACA	169	0.95	0.05	0.00	0.03	0.18
	Anti-Topo	54	0.87	0.13	0.00	0.07	0.27
	Control	84	0.99	0.01	0.00	0.05	
Mantel-Haenszel	SSc	1170					0.22
meta-analysis	Limited	615					0.24
	Diffuse	305					***0.02***
	ACA	573					0.19
	Anti-Topo	286					***0.0001***
	Control	925					

When we combined both discovery and replication cohort we observed a strong association of the *TLR2* Pro631His variant with anti-topoisomerase positive SSc (p<0.00001 OR: 2.55 95%CI:1.85-3.52) (**figure 1**). These results remained significant after Bonferroni correction for multiple testing.

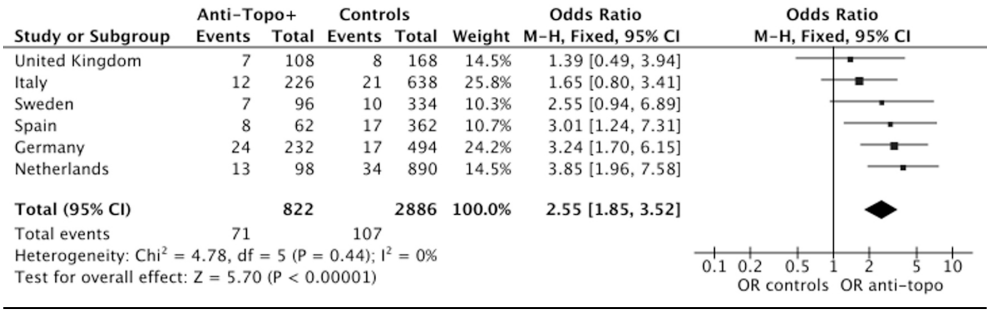


Figure 1. Combined effects of the *TLR2* Pro631His variant on susceptibility for anti-topoisomerase positive SSc (anti-topo). Meta-analysis performed using allele frequencies.

The role of the TLR2 variant in SSc disease severity.

Since we were interested whether this functional polymorphism, which seemed to promote a pro-inflammatory environment could play a role in the severity of SSc, we used a follow-up cohort consisting of 964 SSc patients, which were at least evaluated once a year for complication development. Patients were followed for either 15 years or the time till death, the follow-up period started at the onset of the first symptom after the development of Raynaud's phenomenom. We evaluated the decrease of both FVC and DLCO below 75% of the percentage predicted, the development of pulmonary arterial hypertension measured by right heart catheterization, the development of pulmonary fibrosis by high resolution CT and the disease duration till time of death. Although the TLR2 ligands did not influence patient survival, pulmonary fibrosis or kidney involvement, patients carrying this variant progressed to right heart catheterization proven PAH (total PAH n=40) significantly sooner compared with those patients not carrying the risk allele (Log-Rank p=0.003, Cox proportional hazards ratio: 5.61 (95%CI: 1.53-20.58)) (**figure 2**). There was no significant effect of ATA titers on the development op PAH that otherwise might have biased the observed results.

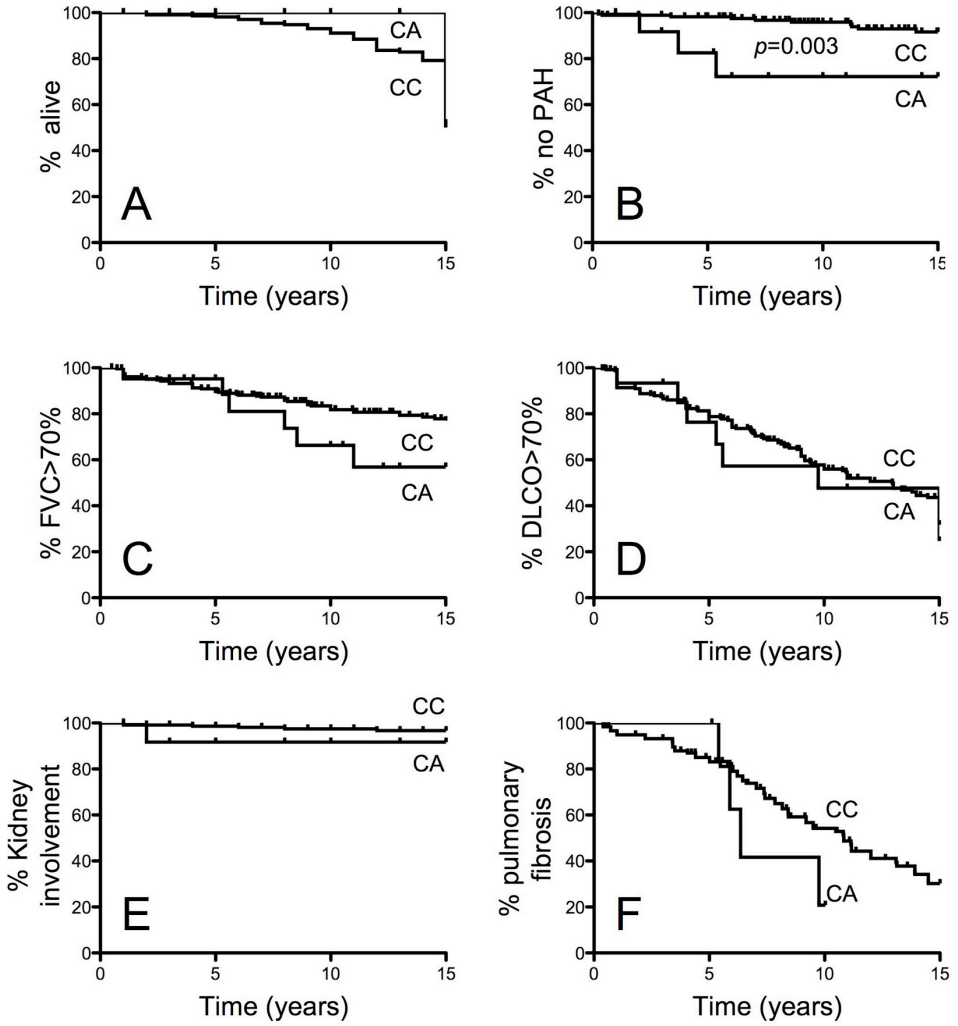


Figure 2. Effect of the *TLR2* Pro631His variant on SSc complication development. No significant effects on death (A), development of either a deterioration of forced vital capacity (FVC) (C), diffusion capacity of the lung for carbon monoxide (DLCO) (D) below 70% of predicted or on the development of pulmonary fibrosis and kidney involvement (E and F) were observed. However, we found an association with the development of pulmonary arterial hypertension (Log-Rank p=0.003, Cox proportional hazards ratio: 5.61 ((95%CI: 1.53-20.58)) (B). Kaplan-Meier graphs were created for 964 patients for a period of 15 years starting at the onset of the first non-Raynaud's symptom.

The TLR2 variant alters the level of cytokine response upon stimulation

Since this polymorphism has previously been shown to influence the expression and function of the TLR2 receptor with regard to antigen uptake, we were interested whether this polymorphism influenced the inflammatory response evoked by the TLR2 ligand Pam3Cys. For this purpose we isolated DCs from 12 patients with and without the rare *TLR2* allele and compared the response upon stimulation. After 24 hrs of stimulation, interestingly, we found that cells from patients carrying the TLR2 variant exhibited a marked increase in production of IL6 (TNFa) upon stimulation with its ligand (both $p<0.0001$) (**figure 3A**). Myeloid DCs from patients revealed similar results as those observed in monocyte-derived DCs. (both $p<0.0001$) (**figure 3B**). No significant differences were present between the current two groups used considering gender, clinical phenotype or medication, controlling for a confounding effect.

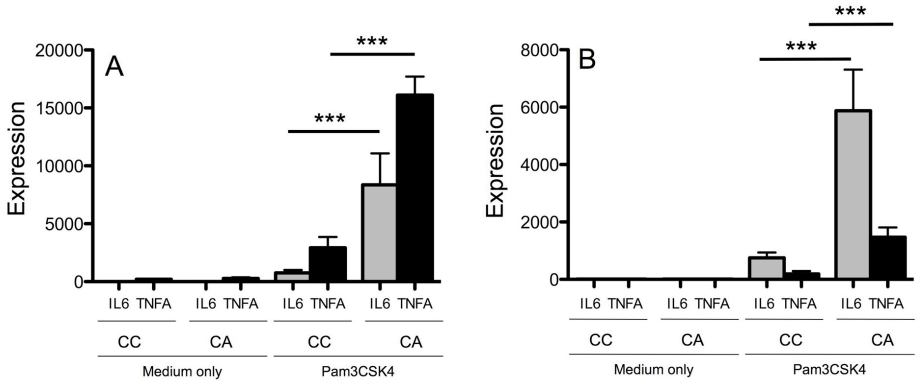


Figure 3. Effects on expression of interleukin 6 (IL6) and tumor necrosis factor alpha (TNFa) in both carriers of the CC (n=9) and CA(n=3) genotypes after stimulation with medium or TLR2 agonist Pam3CSK4 in monocyte derived dendritic cells. The expression of both IL6 and TNFa is significantly increased in carriers of the CA variant (both p<0.0001).

Discussion

Since accumulating evidence suggests a role for TLRs in SSc, we investigated whether genetic variants in five different *TLR* genes influenced SSc susceptibility, severity and cellular function. After an initial study with 3 different ethnic cohorts we found that only one out of fourteen polymorphisms investigated was associated with SSc. More specifically, a rare variant in *TLR2* was found to be associated with SSc patients being positive for anti-topoisomererase antibodies. This association was validated in a large replication cohort consisting of multiple European nationalities. Considering the size and power of the study it is unlikely that the observed effects are coming forth from a type 1 error. A recent genome wide association in SSc, investigated possible associations of tagSNPs with SSc and clinical phenotypes. This study did not investigate possible associations with auto-antibody profiles and did not include this SNP as a haplotype marker, explaining the absence of this SNP coming forth from this effort (13).

The associated polymorphism, *TLR2* Pro631His confers an amino acid change which has previously been functionally investigated. HEK cells transfected with the variant have slowed internal trafficking of TLR2 that not seems to follow conventional pathways. In addition carriers of the variant express more TLR2 variant then wildtype TLR2 on their celmembrane (14). This could explain our observation that dendritic cells patients carrying the variant displayed a more potentiated cytokine response. Although the numbers of donors carrying the minor variant are relatively small compared to the number of donors investigated that do not carry the variant, these results seem to be consistent across monocyte derived dendritic cells and myeloid dendritic cells. Moreover, the production of IL-6 and TNF-a shows little variation within the groups carrying or not carrying the minor TLR2 variant. Interestingly, anti-topo autoantibodies are able to bind to fibroblasts and subsequently attract monocytes in vitro (15). Anti-topo autoantibodies are present in the earliest stages of disease but also in healthy individuals (16,17). A hypothesis would be that anti-topo autoantibodies bind to fibroblasts and attract monocytes. When these monocytes carry the pro-inflammatory *TLR2* variant, both fibroblasts and monocytes might be contributing to a more pronounced inflammatory response and subsequently to a higher risk to develop full-blown

disease. Interestingly, several autologous TLR2 ligands can develop during tissue injury (18,19). In this light it is interesting to note that endothelial cells were shown to express TLR2, which in the case of having the *TLR2* rare variant, might result in an augmented inflammatory response upon stimulation contributing to development of PAH. However, these hypotheses remain to be further investigated.

In conclusion, we show that *TLR2* is involved in SSc phenotype, which might be explained by its altered function upon binding of its ligands. Further research is justified to better determine the precise role of TLR2 in disease onset and or perpetuation. Since, neutralizing antibodies against multiple TLRs are about to enter the clinical arena, more knowledge on the role of TLR2 in SSc may lead to a broadening of the therapeutic armamentarium for this severe condition.

References

1. Gabrielli A, Avvedimento EV, Krieg T. Scleroderma. N Engl J Med 2009; 360(19):1989-2003.

2. Agarwal SK, Reveille JD. The genetics of scleroderma (systemic sclerosis). Curr Opin Rheumatol 2010; 22(2):133-8.

3. Leulier F, Lemaitre B.Toll-like receptors--taking an evolutionary approach.Nat Rev Genet. 2008 Mar;9(3):165-78.

4. Marshak-Rothstein A.Toll-like receptors in systemic autoimmune disease.Nat Rev Immunol. 2006 Nov;6(11):823-35.

5. Roelofs MF, Joosten LA, Abdollahi-Roodsaz S, van Lieshout AW, Sprong T, van den Hoogen FH, van den Berg WB, Radstake TR.The expression of toll-like receptors 3 and 7 in rheumatoid arthritis synovium is increased and costimulation of toll-like receptors 3, 4, and 7/8 results in synergistic cytokine production by dendritic cells.Arthritis Rheum. 2005 Aug; 52(8):2313-22.

6. Van Bon L, Popa C, Huijbens R, Vonk M, York M, Simms R, Hesselstrand R, Wuttge DM, Lafyatis R, Radstake TR. Distinct evolution of TLR-mediated dendritic cell cytokine secretion in patients with limited and diffuse cutaneous systemic sclerosis. Ann Rheum Dis. 2010 Aug;69(8):1539-47. Epub 2010 May 24.

7. Farina GA, York MR, Di Marzio M, Collins CA, Meller S, Homey B, Rifkin IR, Marshak-Rothstein A, Radstake TR, Lafyatis R. Poly(I:C) drives type I IFN- and TGFβ-mediated inflammation and dermal fibrosis simulating altered gene expression in systemic sclerosis. J Invest Dermatol. 2010 Nov;130(11):2583-93. Epub 2010 Jul 8.

8. Veltkamp M, Wijnen PA, van Moorsel CH, Rijkers GT, Ruven HJ, Heron M, Bekers O, Claessen AM, Drent M, van den Bosch JM, Grutters JC. Linkage between Toll-like receptor (TLR) 2 promotor and intron polymorphisms: functional effects and relevance to sarcoidosis. Clin Exp Immunol. 2007 Sep;149(3):453-62. Epub 2007 Jun 12.

9. Assay of locus-specific genetic load implicates rare Toll-like receptor 4 mutations in meningococcal susceptibility. Smirnova I, Mann N, Dols A, Derkx HH, Hibberd ML, Levin M, Beutler B. Proc Natl Acad Sci U S A. 2003 May 13;100(10): 6075-80. Epub 2003 May 2.

10. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. Arthritis Rheum 1980; 23(5): 581-90.

11. LeRoy EC, Black C, Fleischmajer R, Jablonska S, Krieg T, Medsger TA, Jr. et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. J Rheumatol 1988; 15(2):202-5.

12. Pritchard JK, Rosenberg NA. Use of unlinked genetic markers to detect population stratification in association studies. Am J Hum Genet. 1999 Jul; 65 (1):220-8.

13. Radstake TR, Gorlova O, Rueda B, Martin JE, Alizadeh BZ, Palomino-Morales R, Coenen MJ, Vonk MC, Voskuyl AE, Schuerwegh AJ, Broen JC, van Riel PL, van 't Slot R, Italiaander A, Ophoff RA, Riemekasten G, Hunzelmann N, Simeon CP, Ortego-Centeno N, González-Gay MA, González-Escribano MF; Spanish Scleroderma Group, Airo P, van Laar J, Herrick A, Worthington J, Hesselstrand R, Smith V, de Keyser F, Houssiau F, Chee MM, Madhok R, Shiels P, Westhovens R, Kreuter A, Kiener H, de Baere E, Witte T, Padykov L, Klareskog L, Beretta L, Scorza R, Lie BA, Hoffmann-Vold AM, Carreira P, Varga J, Hinchcliff M, Gregersen PK, Lee AT, Ying J, Han Y, Weng SF, Amos CI, Wigley FM, Hummers L, Nelson JL, Agarwal SK, Assassi S, Gourh P, Tan FK, Koeleman BP, Arnett FC, Martin J, Mayes MD. Genome-wide association study of systemic sclerosis identifies CD247 as a new susceptibility locus. Nat Genet. 2010 May;42(5):426-9. Epub 2010 Apr 11.

14. Etokebe GE, Skjeldal F, Nilsen N, Rodionov D, Knezevic J, Bulat-Kardum L, Espevik T, Bakke O, Dembic Z. Toll-like receptor 2 (P631H) mutant impairs membrane internalization and is a dominant negative allele. Scand J Immunol. 2010 May;71(5):369-81.

15. Hénault J, Robitaille G, Senécal JL, Raymond Y. DNA topoisomerase I binding to fibroblasts induces monocyte adhesion and activation in the presence of anti-topoisomerase I autoantibodies from systemic sclerosis patients. Arthritis Rheum. 2006 Mar;54(3):963-73.

16. Kallenberg CG, Wouda AA, Hoet MH, van Venrooij WJ. Development of connective tissue disease in patients presenting with Raynaud's phenomenon: a six year follow up with emphasis on the predictive value of antinuclear antibodies as detected by immunoblotting. Ann Rheum Dis. 1988 Aug;47(8):634-41.

17. Weiner ES, Hildebrandt S, Senécal JL, Daniels L, Noell S, Joyal F, Roussin A, Earnshaw W, Rothfield NF. Prognostic significance of anticentromere antibodies and anti-topoisomerase I antibodies in Raynaud's disease. A prospective study. Arthritis Rheum. 1991 Jan;34(1):68-77.

18. Scheibner KA, Lutz MA, Boodoo S, Fenton MJ, Powell JD, Horton MR. Hyaluronan fragments act as an endogenous danger signal by engaging TLR2. J Immunol. 2006 Jul 15;177(2): 1272-81.

19. Schaefer L, Babelova A, Kiss E, Hausser HJ, Baliova M, Krzyzankova M, Marsche G, Young MF, Mihalik D, Götte M, Malle E, Schaefer RM, Gröne HJ. The matrix component biglycan is proinflammatory and signals through Toll-like receptors 4 and 2 in macrophages. J Clin Invest. 2005 Aug;115(8):2223-33. Epub 2005 Jul 14

20. Kastelijn EA, van Moorsel CH, Rijkers GT, Ruven HJ, Karthaus V, Kwakkel-van Erp JM, van de Graaf EA, Zanen P, van Kessel DA, Grutters JC, van den Bosch JM. Polymorphisms in innate immunity genes associated with development of bronchiolitis obliterans after lung transplantation. J Heart Lung Transplant. 2010 Jun;29(6):665-71. Epub 2010 Mar 15.

21. Tsui FW, Xi N, Rohekar S, Riarh R, Bilotta R, Tsui HW, Inman RD. Toll-like receptor 2 variants are associated with acute reactive arthritis. Arthritis Rheum. 2008 Nov;58(11):3436-8.

22. Henckaerts L, Pierik M, Joossens M, Ferrante M, Rutgeerts P, Vermeire S. Mutations in pattern recognition receptor genes modulate seroreactivity to microbial antigens in patients with inflammatory bowel disease. Gut. 2007 Nov;56(11):1536-42. Epub 2007 Jun 26.

23. Saçkesen C, Karaaslan C, Keskin O, Tokol N, Tahan F, Civelek E, Soyer OU, Adalioglu G, Tuncer A, Birben E, Oner C, Kalayci O. The effect of polymorphisms at the CD14 promoter and the TLR4 gene on asthma phenotypes in Turkish children with asthma. Allergy. 2005 Dec;60(12): 1485-92.

24. Huang H, Shiffman ML, Friedman S, Venkatesh R, Bzowej N, Abar OT, Rowland CM, Catanese JJ, Leong DU, Sninsky JJ, Layden TJ, Wright TL, White T, Cheung RC. A 7 gene signature identifies the risk of developing cirrhosis in patients with chronic hepatitis C. Hepatology. 2007 Aug;46(2):297-306.

25. Enevold C, Oturai AB, Sørensen PS, Ryder LP, Koch-Henriksen N, Bendtzen K. Polymorphisms of innate pattern recognition receptors, response to interferon-beta and development of neutralizing antibodies in multiple sclerosis patients. Mult Scler. 2010 Aug; 16(8):942-9. Epub 2010 Jul 1.

26. Shen N, Fu Q, Deng Y, Qian X, Zhao J, Kaufman KM, Wu YL, Yu CY, Tang Y, Chen JY, Yang W, Wong M, Kawasaki A, Tsuchiya N, Sumida T, Kawaguchi Y, Howe HS, Mok MY, Bang SY, Liu FL, Chang DM, Takasaki Y, Hashimoto H, Harley JB, Guthridge JM, Grossman JM, Cantor RM, Song YW, Bae SC, Chen S, Hahn BH, Lau YL, Tsao BP. Sex-specific association of X-linked Toll-like receptor 7 (TLR7) with male systemic lupus erythematosus. Proc Natl Acad Sci U S A. 2010 Sep 7;107(36):15838-43. Epub 2010 Aug 23.

27. Møller-Larsen S, Nyegaard M, Haagerup A, Vestbo J, Kruse TA, Børglum AD. Association analysis identifies TLR7 and TLR8 as novel risk genes in asthma and related disorders. Thorax. 2008 Dec; 63(12):1064-9. Epub 2008 Aug 5.

28. Schott E, Witt H, Neumann K, Bergk A, Halangk J, Weich V, Müller T, Puhl G, Wiedenmann B, Berg T. Association of

TLR7 single nucleotide polymorphisms with chronic HCV-infection and response to interferon-a-based therapy. J Viral Hepat. 2008 Jan;15(1):71-8.

29. Georgel P, Macquin C, Bahram S. The heterogeneous allelic repertoire of human toll-like receptor (TLR) genes. PLoS One. 2009 Nov 17;4(11):e7803.

30. Engin A, Arslan S, Kizildag S, Oztürk H, Elaldi N, Dökmetas I, Bakir M. Toll-like receptor 8 and 9 polymorphisms in Crimean-Congo hemorrhagic fever. Microbes Infect. 2010 Nov;12(12-13): 1071-8. Epub 2010 Jul 30.

31. Gantier MP, Irving AT, Kaparakis-Liaskos M, Xu D, Evans VA, Cameron PU, Bourne JA, Ferrero RL, John M, Behlke MA, Williams BR. Genetic modulation of TLR8 response following bacterial phagocytosis. Hum Mutat. 2010 Sep; 31(9):1069-79.

32. Davila S, Hibberd ML, Hari Dass R, Wong HE, Sahiratmadja E, Bonnard C, Alisjahbana B, Szeszko JS, Balabanova Y, Drobniewski F, van Crevel R, van de Vosse E, Nejentsev S, Ottenhoff TH, Seielstad M. Genetic association and expression studies indicate a role of toll-like receptor 8 in pulmonary tuberculosis. PLoS Genet. 2008 Oct;4(10):e1000218. Epub 2008 Oct 10.

33. Enevold C, Radstake TR, Coenen MJ, Fransen J, Toonen EJ, Bendtzen K, van Riel PL. Multiplex screening of 22 single-nucleotide polymorphisms in 7 Toll-like receptors: an association study in rheumatoid arthritis. J Rheumatol. 2010 May;37(5):905-10. Epub 2010 Mar 1.

34. Liao WL, Chen RH, Lin HJ, Liu YH, Chen WC, Tsai Y, Wan L, Tsai FJ. Toll-like receptor gene polymorphisms are associated with susceptibility to Graves' ophthalmopathy in Taiwan males. BMC Med Genet. 2010 Nov 5;11:154.

Supplementary Table 1. The clinical characteristics of the 6 SSc cohorts included in this study.

Population	The Netherlands	Spain	Germany	Sweden	Italy	United Kingdom
Number	323	188	312	193	337	269
Age (years (SD))	59 (15)	58 (13)	57(12)	53 (14)	55(13)	54 (12)
Disease duration (months (SD))	137 (84)	144 (90)	113 (109)	82 (73)	140 (138)	155 (92)
Female %	79.0	81.8	88.4	76.9	82.3	81.2
Limited phenotype %	74.6	72.2	59.5	82.4	73.3	74.2
Positivity anti-topo %	23.2	26.8	26.7	16.9	32.9	12.4
Positivity ACA %	32.2	46.7	41.4	37.1	32.1	37.3

Supplementary Table 2: Overview of the genotype and allele frequencies of *TLR2* rs1898830 in the individual populations.

cohort	phenotype	n	GG	GA	AA	p	G	p
Italy	**SSc**	170	0.16	0.35	0.48	0.95	0.34	0.87
	lcSSc	121	0.15	0.35	0.50	0.96	0.32	0.79
	dcSSc	44	0.18	0.36	0.45	0.85	0.36	0.62
	ACA	64	0.19	0.34	0.47	0.79	0.36	0.59
	Anti-Topo	79	0.13	0.33	0.54	0.73	0.29	0.36
	Controls	88	0.16	0.36	0.48		0.34	
The Netherlands	**SSc**	82	0.07	0.41	0.52	0.75	0.27	0.79
	lcSSc	62	0.13	0.40	0.47	0.61	0.33	0.36
	dcSSc	18	0.17	0.39	0.44	0.54	0.36	0.34
	ACA	57	0.12	0.39	0.49	0.79	0.32	0.55
	Anti-Topo	21	0.05	0.38	0.57	0.94	0.24	0.59
	Controls	252	0.09	0.38	0.52		0.28	
Spain	**SSc**	177	0.13	0.38	0.49	0.41	0.32	0.87
	lcSSc	128	0.13	0.39	0.48	0.68	0.32	0.86
	dcSSc	48	0.15	0.33	0.52	0.34	0.31	0.99
	ACA	73	0.15	0.34	0.51	0.27	0.32	0.83
	Anti-Topo	29	0.07	0.38	0.55	0.76	0.26	0.45
	Controls	179	0.09	0.44	0.47		0.31	
Mantel-Haenszel	**SSc**	429						0.92
meta-analysis	**lcSSc**	311						0.68
	dcSSc	110						0.46
	ACA	194						0.44
	Anti-Topo	129						0.16
	Controls	519						

Supplementary Table 3: Overview of the genotype and allele frequencies of *TLR4* rs7873784 in the individual populations

cohort	phenotype	n	CC	CG	GG	p	C	p
Italy	**SSc**	175	0.77	0.22	0.01	0.155	0.88	0.21
	lcSSc	119	0.80	0.19	0.01	0.23	0.89	0.13
	dcSSc	44	0.80	0.20	0.00	0.37	0.90	0.26
	ACA	63	0.78	0.21	0.02	0.65	0.88	0.4
	Anti-Topo	78	0.77	0.23	0.00	0.31	0.88	0.27
	Controls	85	0.72	0.25	0.04		0.84	
Netherlands	**SSc**	81	0.68	0.27	0.05	0.39	0.81	0.21
	lcSSc	63	0.67	0.29	0.05	0.41	0.81	0.21
	dcSSc	17	0.71	0.24	0.06	0.71	0.82	0.61
	ACA	23	0.57	0.39	0.04	0.14	0.76	0.09
	Anti-Topo	19	0.84	0.16	0.00	0.6	0.92	0.34
	Controls	239	0.74	0.23	0.03		0.86	
Spain	**SSc**	170	0.74	0.26	0.01	0.22	0.86	0.4
	lcSSc	121	0.73	0.26	0.01	0.38	0.86	0.6
	dcSSc	48	0.75	0.25	0.00	0.44	0.88	0.43
	ACA	71	0.72	0.28	0.00	0.37	0.86	0.68
	Anti-Topo	28	0.68	0.32	0.00	0.59	0.84	0.99
	Controls	184	0.71	0.26	0.03		0.86	
Mantel-Haenszel	**SSc**	426						0.58
meta-analysis	**lcSSc**	303						0.66
	dcSSc	109						0.34
	ACA	157						0.95
	Anti-Topo	125						0.25
	Controls	508						

Supplementary Table 4: Overview of the genotype and allele frequencies of *TLR4* rs4986791 in the individual populations

cohort	phenotype	n	TT	TC	CC	p	T	p
Italy	**SSc**	169	0.00	0.13	0.87	0.075	0.07	0.08
	lcSSc	114	0.00	0.12	0.88	0.077	0.06	0.09
	dcSSc	43	0.00	0.19	0.81	0.013	0.09	0.03
	ACA	62	0.00	0.18	0.82	0.025	0.09	0.03
	Anti-Topo	74	0.00	0.08	0.92	0.52	0.04	0.53
	Controls	82	0.00	0.05	0.95		0.02	
Netherlands	**SSc**	82	0.00	0.12	0.88	0.99	0.06	0.99
	lcSSc	64	0.00	0.09	0.91	0.6	0.05	0.67
	dcSSc	17	0.00	0.24	0.76	0.155	0.12	0.15
	ACA	23	0.00	0.09	0.91	0.49	0.04	0.67
	Anti-Topo	20	0.00	0.20	0.80	0.29	0.10	0.3
	Controls	239	0.00	0.12	0.88		0.06	
Spain	**SSc**	179	0.01	0.09	0.90	0.86	0.05	0.73
	lcSSc	128	0.00	0.09	0.91	0.93	0.05	0.99
	dcSSc	50	0.02	0.10	0.88	0.15	0.07	0.31
	ACA	74	0.00	0.09	0.91	0.55	0.05	0.93
	Anti-Topo	29	0.03	0.03	0.93	0.08	0.05	0.99
	Controls	187	0.00	0.09	0.91		0.05	
Mantel-Haenszel	**SSc**	430						0.21
meta-analysis	**lcSSc**	306						0.61
	dcSSc	110						0.01
	ACA	159						0.26
	Anti-Topo	123						0.32
	Controls	508						

Supplementary Table 5: Overview of the genotype and allele frequencies of *TLR4* rs4986790 in the individual populations

cohort	phenotype	n	GG	GA	AA	p	G	p
Italy	**SSc**	172	0.00	0.07	0.93	0.149	0.03	0.16
	lcSSc	120	0.00	0.12	0.88	0.26	0.06	0.36
	dcSSc	43	0.00	0.21	0.79	0.02	0.10	0.04
	ACA	86	0.00	0.07	0.93	0.09	0.03	0.12
	Anti-Topo	77	0.00	0.09	0.91	0.77	0.05	0.78
	Controls	87	0.00	0.13	0.87		0.07	
Netherlands	**SSc**	82	0.01	0.12	0.87	0.99	0.07	0.59
	lcSSc	64	0.02	0.09	0.89	0.54	0.06	0.99
	dcSSc	17	0.00	0.24	0.76	0.34	0.12	0.27
	ACA	23	0.00	0.09	0.91	0.87	0.04	0.99
	Anti-Topo	20	0.00	0.20	0.80	0.34	0.10	0.5
	Controls	243	0.00	0.12	0.88		0.06	
Spain	**SSc**	181	0.01	0.10	0.90	0.79	0.06	0.61
	lcSSc	129	0.00	0.10	0.90	0.77	0.05	0.85
	dcSSc	51	0.02	0.10	0.88	0.16	0.07	0.44
	ACA	73	0.00	0.10	0.90	0.9	0.05	0.99
	Anti-Topo	31	0.03	0.06	0.90	0.12	0.06	0.52
	Controls	187	0.00	0.09	0.91		0.05	
Mantel-Haenszel	**SSc**	435						0.67
meta-analysis	**lcSSc**	313						0.45
	dcSSc	111						0.02
	ACA	182						0.51
	Anti-Topo	128						0.32
	Controls	517						

Supplementary Table 6: Overview of the genotype and allele frequencies of *TLR7* rs3853839 in the individual populations

cohort	phenotype	n	GG	GC	CC	p	G	p
Italy	**SSc**	142	0.04	0.31	0.65	0.84	0.19	0.61
	lcSSc	97	0.04	0.36	0.60	0.79	0.22	0.99
	dcSSc	35	0.05	0.20	0.75	0.4	0.15	0.26
	ACA	51	0.05	0.39	0.56	0.7	0.24	0.59
	Anti-Topo	63	0.04	0.28	0.68	0.71	0.18	0.44
	Controls	69	0.06	0.31	0.63		0.22	
Netherlands	**SSc**	66	0.01	0.18	0.80	0.95	0.10	0.63
	lcSSc	52	0.02	0.18	0.80	0.93	0.11	0.99
	dcSSc	13	0.00	0.19	0.81	0.99	0.09	0.75
	ACA	18	0.04	0.13	0.83	0.37	0.11	0.99
	Anti-Topo	16	0.00	0.20	0.80	0.91	0.10	0.78
	Controls	133	0.02	0.20	0.77		0.13	
Spain	**SSc**	140	0.09	0.24	0.67	0.87	0.21	0.84
	lcSSc	98	0.09	0.21	0.70	0.94	0.20	0.65
	dcSSc	41	0.06	0.31	0.63	0.38	0.22	0.99
	ACA	55	0.10	0.14	0.75	0.45	0.17	0.41
	Anti-Topo	25	0.10	0.26	0.65	0.94	0.23	0.85
	Controls	146	0.10	0.22	0.67		0.22	
Mantel-Haenszel	**SSc**	347						0.46
meta-analysis	**lcSSc**	247						0.71
	dcSSc	89						0.35
	ACA	125						0.63
	Anti-Topo	104						0.46
	Controls	348						

Supplementary Table 7: Overview of the genotype and allele frequencies of TLR7 rs179008 in the individual populations (females only)

cohort	phenotype	n	TT	TA	AA	p	T	p
Italy	**SSc**	139	0.03	0.25	0.71	0.89	0.16	0.58
	lcSSc	95	0.03	0.27	0.71	0.78	0.16	0.55
	dcSSc	34	0.05	0.21	0.74	0.75	0.15	0.56
	ACA	50	0.02	0.34	0.65	0.6	0.19	0.99
	Anti-Topo	63	0.05	0.20	0.75	0.77	0.15	0.51
	Controls	68	0.05	0.27	0.68		0.18	
Netherlands	**SSc**	66	0.08	0.35	0.57	0.86	0.26	0.73
	lcSSc	52	0.09	0.37	0.54	0.67	0.28	0.45
	dcSSc	14	0.06	0.29	0.65	0.91	0.21	0.99
	ACA	18	0.04	0.43	0.52	0.86	0.26	0.69
	Anti-Topo	16	0.10	0.15	0.75	0.15	0.18	0.67
	Controls	204	0.07	0.33	0.59		0.24	
Spain	**SSc**	142	0.06	0.29	0.66	0.98	0.20	0.92
	lcSSc	102	0.07	0.31	0.62	0.7	0.22	0.43
	dcSSc	40	0.02	0.24	0.74	0.72	0.14	0.33
	ACA	58	0.04	0.32	0.64	0.96	0.20	0.99
	Anti-Topo	23	0.03	0.31	0.66	0.99	0.19	0.99
	Controls	146	0.05	0.29	0.66		0.19	
Mantel-Haenszel	**SSc**	348						0.99
meta-analysis	**lcSSc**	249						0.48
	dcSSc	88						0.19
	ACA	126						0.84
	Anti-Topo	102						0.4
	Controls	418						

Supplementary Table 8: Overview of the genotype and allele frequencies of *TLR7* rs2302267 in the individual populations (females only)

cohort	phenotype	n	TT	TG	GG	p	T	p
Italy	**SSc**	139	0.84	0.16	0.01	0.8	0.92	0.72
	lcSSc	94	0.83	0.16	0.01	0.99	0.91	0.87
	dcSSc	35	0.84	0.16	0.00	0.27	0.92	0.99
	ACA	50	0.87	0.11	0.02	0.51	0.93	0.83
	Anti-Topo	63	0.80	0.20	0.00	0.6	0.90	0.57
	Controls	70	0.82	0.18	0.00		0.91	
Netherlands	**SSc**	67	0.94	0.04	0.03	0.48	0.95	0.18
	lcSSc	52	0.89	0.05	0.06	0.17	0.92	0.81
	dcSSc	14	0.89	0.06	0.06	0.59	0.92	0.99
	ACA	18	0.91	0.00	0.09	0.77	0.91	0.99
	Anti-Topo	16	0.90	0.10	0.00	0.31	0.95	0.99
	Controls	202	0.90	0.04	0.06		0.92	
Spain	**SSc**	141	0.93	0.06	0.01	0.22	0.96	0.05
	lcSSc	101	0.94	0.06	0.01	0.43	0.96	0.08
	dcSSc	39	0.92	0.08	0.00	0.51	0.96	0.32
	ACA	56	0.94	0.06	0.00	0.43	0.97	0.1
	Anti-Topo	24	0.90	0.07	0.03	0.86	0.93	0.99
	Controls	149	0.90	0.06	0.04		0.93	
Mantel-Haenszel	**SSc**	347						0.09
meta-analysis	**lcSSc**	247						0.55
	dcSSc	89						0.49
	ACA	124						0.19
	Anti-Topo	103						0.78
	Controls	421						

Supplementary Table 9: Overview of the genotype and allele frequencies of *TLR7* rs5743781 in the individual populations (females only)

cohort	phenotype	n	CC	CT	TT	p	C	p
Italy	**SSc**	138	0.00	0.01	0.99	0.99	0.00	0.99
	lcSSc	94	0.00	0.01	0.99	0.83	0.00	0.99
	dcSSc	34	0.00	0.00	1.00	0.48	0.00	0.99
	ACA	89	0.00	0.01	0.99	0.4	0.00	0.99
	Anti-Topo	63	0.00	0.01	0.99	0.99	0.01	0.99
	Controls	70	0.00	0.01	0.99		0.01	
Netherlands	**SSc**	66	0.00	0.00	1.00	N.A.	0.00	N.A.
	lcSSc	51	0.00	0.00	1.00	N.A.	0.00	N.A.
	dcSSc	14	0.00	0.00	1.00	N.A.	0.00	N.A.
	ACA	46	0.00	0.00	1.00	N.A.	0.00	N.A.
	Anti-Topo	15	0.00	0.00	1.00	N.A.	0.00	N.A.
	Controls	192	0.00	0.00	1.00		0.00	
Spain	**SSc**	149	0.01	0.02	0.98	0.86	0.01	0.99
	lcSSc	102	0.00	0.02	0.98	0.71	0.01	0.71
	dcSSc	41	0.00	0.04	0.96	0.52	0.02	0.99
	ACA	58	0.00	0.00	1.00	0.45	0.00	0.33
	Anti-Topo	25	0.00	0.03	0.97	0.99	0.02	0.99
	Controls	144	0.00	0.02	0.98		0.01	
Mantel-Haenszel	**SSc**	353						0.88
meta-analysis	**lcSSc**	248						0.73
	dcSSc	89						0.78
	ACA	194						0.24
	Anti-Topo	103						0.78
	Controls	406						

Supplementary Table 10: Overview of the genotype and allele frequencies of *TLR8* rs3764879 in the individual populations (females only)

cohort	phenotype	n	CC	CG	GG	p	C	p
Italy	**SSc**	142	0.08	0.37	0.55	0.78	0.27	0.73
	lcSSc	96	0.10	0.40	0.50	0.92	0.30	0.81
	dcSSc	35	0.07	0.27	0.66	0.32	0.20	0.19
	ACA	51	0.13	0.42	0.45	0.75	0.34	0.48
	Anti-Topo	63	0.06	0.35	0.58	0.62	0.24	0.41
	Controls	71	0.08	0.42	0.51		0.29	
Netherlands	**SSc**	67	0.06	0.29	0.65	0.75	0.20	0.7
	lcSSc	52	0.05	0.26	0.69	0.96	0.18	0.89
	dcSSc	14	0.12	0.41	0.47	0.06	0.32	0.08
	ACA	18	0.04	0.17	0.78	0.45	0.13	0.65
	Anti-Topo	16	0.05	0.25	0.70	0.79	0.18	0.99
	Controls	198	0.04	0.29	0.67		0.18	
Spain	**SSc**	138	0.15	0.33	0.52	0.65	0.32	0.41
	lcSSc	98	0.14	0.35	0.51	0.74	0.31	0.42
	dcSSc	39	0.16	0.29	0.55	0.73	0.31	0.67
	ACA	58	0.13	0.29	0.58	0.9	0.27	0.9
	Anti-Topo	23	0.07	0.34	0.59	0.99	0.24	0.6
	Controls	146	0.12	0.33	0.55		0.28	
Mantel-Haenszel	**SSc**	347						0.55
meta-analysis	**lcSSc**	246						0.56
	dcSSc	88						0.79
	ACA	127						0.98
	Anti-Topo	102						0.34
	Controls	414						

Supplementary Table 11: Overview of the genotype and allele frequencies of *TLR8* rs3764880 in the individual populations (females only)

cohort	phenotype	n	AA	AG	GG	p	A	p
Italy	**SSc**	141	0.09	0.36	0.55	0.99	0.27	0.99
	lcSSc	96	0.10	0.40	0.50	0.82	0.30	0.46
	dcSSc	35	0.07	0.27	0.66	0.68	0.20	0.39
	ACA	51	0.13	0.42	0.45	0.53	0.34	0.25
	Anti-Topo	62	0.06	0.35	0.59	0.89	0.24	0.78
	Controls	67	0.08	0.36	0.56		0.26	
Netherlands	**SSc**	65	0.06	0.30	0.64	0.92	0.21	0.72
	lcSSc	50	0.05	0.27	0.68	0.63	0.18	0.34
	dcSSc	14	0.17	0.39	0.44	0.27	0.36	0.16
	ACA	18	0.04	0.17	0.78	0.54	0.13	0.3
	Anti-Topo	15	0.05	0.26	0.68	0.99	0.18	0.82
	Controls	194	0.08	0.29	0.63		0.23	
Spain	**SSc**	138	0.17	0.30	0.53	0.91	0.32	0.99
	lcSSc	101	0.15	0.33	0.52	0.7	0.31	0.92
	dcSSc	37	0.20	0.24	0.57	0.91	0.32	0.99
	ACA	58	0.12	0.25	0.63	0.42	0.25	0.19
	Anti-Topo	22	0.11	0.29	0.61	0.61	0.25	0.39
	Controls	144	0.18	0.28	0.54		0.32	
Mantel-Haenszel	**SSc**	344						0.87
meta-analysis	**lcSSc**	247						0.88
	dcSSc	86						0.94
	ACA	128						0.41
	Anti-Topo	100						0.3
	Controls	406						

Supplementary Table 12: Overview of the genotype and allele frequencies of *TLR8* rs5741883 in the individual populations (females only)

cohort	phenotype	n	TT	TC	CC	p	T	p
Italy	**SSc**	141	0.09	0.29	0.62	0.76	0.24	0.38
	lcSSc	96	0.07	0.28	0.66	0.96	0.20	0.89
	dcSSc	35	0.14	0.30	0.57	0.41	0.28	0.16
	ACA	51	0.05	0.27	0.69	0.82	0.18	0.87
	Anti-Topo	63	0.10	0.34	0.56	0.4	0.27	0.15
	Controls	70	0.07	0.25	0.68		0.20	
Netherlands	**SSc**	54	0.16	0.26	0.57	0.05	0.29	0.09
	lcSSc	42	0.17	0.26	0.57	0.05	0.30	0.07
	dcSSc	11	0.14	0.21	0.64	0.68	0.25	0.6
	ACA	15	0.16	0.26	0.58	0.29	0.29	0.37
	Anti-Topo	12	0.13	0.27	0.60	0.24	0.27	0.62
	Controls	200	0.15	0.12	0.73		0.21	
Spain	**SSc**	142	0.08	0.30	0.62	0.99	0.23	0.99
	lcSSc	102	0.08	0.31	0.61	0.99	0.23	0.99
	dcSSc	39	0.08	0.29	0.63	0.99	0.22	0.99
	ACA	57	0.10	0.32	0.58	0.79	0.26	0.52
	Anti-Topo	23	0.10	0.24	0.66	0.94	0.22	0.99
	Controls	154	0.08	0.30	0.62		0.23	
Mantel-Haenszel	**SSc**	337						0.18
meta-analysis	**lcSSc**	240						0.3
	dcSSc	86						0.3
	ACA	123						0.49
	Anti-Topo	98						0.27
	Controls	424						

Supplementary Table 13: Overview of the genotype and allele frequencies of *TLR8* rs5744088 in the individual populations (females only)

cohort	phenotype	n	GG	GC	CC	p	G	p
Italy	**SSc**	142	0.79	0.18	0.03	0.73	0.88	0.36
	lcSSc	97	0.80	0.17	0.02	0.51	0.89	0.33
	dcSSc	35	0.80	0.18	0.02	0.92	0.89	0.53
	ACA	51	0.86	0.11	0.03	0.35	0.91	0.17
	Anti-Topo	63	0.75	0.22	0.04	0.99	0.85	0.87
	Controls	71	0.74	0.21	0.04		0.85	
Netherlands	**SSc**	66	0.70	0.19	0.11	0.59	0.80	0.36
	lcSSc	51	0.73	0.16	0.11	0.43	0.81	0.66
	dcSSc	14	0.67	0.22	0.11	0.46	0.78	0.6
	ACA	18	0.70	0.26	0.04	0.83	0.83	0.99
	Anti-Topo	16	0.75	0.15	0.10	0.48	0.83	0.81
	Controls	206	0.74	0.19	0.07		0.83	
Spain	**SSc**	142	0.76	0.20	0.04	0.86	0.86	0.99
	lcSSc	102	0.72	0.23	0.06	0.55	0.83	0.48
	dcSSc	40	0.88	0.12	0.00	0.28	0.94	0.08
	ACA	58	0.71	0.25	0.04	0.56	0.84	0.99
	Anti-Topo	25	0.77	0.13	0.10	0.71	0.84	0.67
	Controls	150	0.78	0.17	0.05		0.86	
Mantel-Haenszel	**SSc**	350						0.98
meta-analysis	**lcSSc**	250						0.8
	dcSSc	90						0.21
	ACA	128						0.33
	Anti-Topo	104						0.87
	Controls	426						

Supplementary Table 14: Overview of the genotype and allele frequencies of *TLR9* rs187084 in the individual populations

cohort	phenotype	n	CC	CT	TT	p	C	p
Italy	**SSc**	176	0.18	0.47	0.35	0.48	0.42	0.45
	lcSSc	120	0.18	0.53	0.30	0.93	0.44	0.76
	dcSSc	44	0.18	0.39	0.43	0.17	0.38	0.24
	ACA	65	0.20	0.46	0.34	0.61	0.43	0.73
	Anti-Topo	77	0.18	0.45	0.36	0.45	0.41	0.44
	Controls	87	0.18	0.54	0.28		0.45	
Netherlands	**SSc**	83	0.13	0.51	0.36	0.49	0.39	0.28
	lcSSc	64	0.13	0.52	0.36	0.53	0.38	0.32
	dcSSc	18	0.17	0.44	0.39	0.68	0.39	0.61
	ACA	23	0.17	0.48	0.35	0.84	0.41	0.88
	Anti-Topo	19	0.05	0.58	0.37	0.42	0.34	0.31
	Controls	247	0.17	0.54	0.30		0.44	
Spain	**SSc**	183	0.19	0.45	0.36	0.89	0.42	0.65
	lcSSc	129	0.16	0.45	0.40	0.66	0.38	0.41
	dcSSc	50	0.22	0.42	0.36	0.89	0.43	0.82
	ACA	73	0.18	0.42	0.40	0.86	0.39	0.62
	Anti-Topo	31	0.19	0.45	0.35	0.99	0.42	0.99
	Controls	180	0.18	0.44	0.38		0.40	
Mantel-Haenszel	**SSc**	442						0.18
meta-analysis	**lcSSc**	313						0.18
	dcSSc	112						0.48
	ACA	161						0.52
	Anti-Topo	127						0.33
	Controls	514						

Chapter 8

Variants of *Pre-B-cell colony-enhancing factor* predispose to Systemic Sclerosis and pulmonary arterial hypertension development

Broen JCA*, Gourh P*, Vonk MC*, Beretta L, Niederer F, Rueda B, Geurts-van Bon L, Brouwer C, Hesselstrand R, Herrick A, Worthington J, Hunzelman N, Denton C, Fonseca, Riemekasten G, Kiener H, Scorza R, Simeon CP, Fonollosa V (for the Spanish Systemic Sclerosis group), Carreira P, Ortego-Centeno N, Gonzalez-Gay MA,[15] Airo' P,[16] Coenen MJH,[17] Mayes M,[2] Kyburz D, Arnett F.C*, Martin J* and Radstake TRDJ*

Submitted

Abstract

Aim: To further elucidate the role of Pre-B-cell colony-enhancing factor (PBEF) in systemic sclerosis (SSc) related pulmonary complications.

Methods: We measured PBEF serum expression levels in serum of patients with and without pulmonary complications by ELISA. Next, these levels were correlated with previously described functional polymorphisms genotyped using custom Taqman 5´allelic discrimination assays. Furthermore, the effect of these polymorphisms on development of pulmonary arterial hypertension (PAH) in time was assessed using a follow up cohort consisting of 964 SSc patients. Next, we were interested in the effect of these polymorphisms on SSc development in general and we genotyped DNA from 2737 SSc patients and 1913 matched healthy controls, both from 8 different ethnic populations.

Results: PBEF levels were significantly higher in patients with PAH compared to patients without, and to patients with other pulmonary complications (P=0.02). Next we found that two polymorphisms in the promoter region of *PBEF* influence protein expression levels, having no minor alleles at both loci resulted in a significantly higher PBEF level in serum (p<0.001) Carrying of no minor variants at these loci was associated with less PAH development during 15 years of follow up compared to carriers of minor alleles (Log Rank p=0.02, HR 2.38, 95%CI: 1.2-5.3). Interestingly, the combined *PBEF* -1543CC -1001TT genotype, hence carrying no minor alleles, was found associated with SSc susceptibility (p=0.009 OR 1.20 (95% CI 1.05-1.37).

Conclusion: Our data identify *PBEF* as a novel candidate gene that influences the development of PAH and SSc susceptibility.

Introduction

Systemic sclerosis (SSc) is a connective tissue disease in which patients suffer from extensive fibrosis of the skin and internal organs. Furthermore, SSc is characterized by endothelial cell damage and immune dysfunction (1). Among multiple other disease manifestations, the pulmonary complications are often regarded as the most severe and life threatening. Frequently encountered complications are pulmonary arterial hypertension (PAH), pulmonary fibrosis and decreasing forced vital capacity (FVC) and diffusion capacity of the lungs for carbonmonoxide (DLCO). A recent microarray analysis of SSc patients with PAH identified Pre B-cell colony-enhancing factor (PBEF) as a potentially interesting molecule that was increasingly expressed in peripheral blood leucocytes of PAH positive patients (2). PBEF has been associated with pulmonary pathology numerous times. For instance, PBEF is able to mount the production of pro-inflammatory cytokines such as IL1β, IL6, IL10 and TNFα by human monocytes and CXCL8 expression by pulmonary epithelial cells (3). This spectrum of mediators is increased in SSc patients (4). In addition, it is postulated that PBEF plays an important role in pulmonary homeostasis. PBEF is upregulated by mechanical stretch and induces pulmonary artery permeability (5-8). Genetic variants in *PBEF* have been found to mediate the extent of lung inflammation and survival in acute lung injury (7,9).

Taken together, it is tempting to speculate that PBEF is involved in PAH development as observed in SSc. For this reason we were interested to validate the observation of increased PBEF levels in PAH and investigated two promoter polymorphisms able to alter *PBEF* expression levels for their contribution to in vivo PBEF levels, the development of PAH throughout time and SSc susceptibility overall.

Methods

Patients and Controls

All included patients fulfilled the 1980 American College of Rheumatology (ACR) classification criteria for SSc (10). The local ethical committee from each center approved the study. Both patients and controls were included in the study after written informed consent. All patients included in this study were classified as having limited cutaneous (lcSSc) or diffuse cutaneous SSc (dcSSc) using the criteria postulated by LeRoy (11). Controls were derived from the same country and region as the patients were derived from. Controls have been extensively tested for population stratification on a genome wide basis previously (12, 13). The presence of pulmonary fibrosis was investigated by a high resolution computed tomography scan. Restrictive syndrome and diffusion capacity of the lungs was defined as a forced vital capacity (FVC) < 70% of the predicted value and a diffusion capacity of the lung for carbon monoxide (DLCO) of less than 70% of predicted. Pulmonary arterial hypertension was diagnosed by right heart catheterization and considered confirmed if the mean pulmonary artery pressure was greater than 25 mmHg at rest with a normal wedge pressure (<15mmHg). For the measurement of PBEF in serum we included 40 Dutch SSc patients. Furthermore, we used follow up data on PAH development, FVC and DLCO decline, and development of pulmonary fibrosis from 964 Dutch, Swedish and Italian patients, starting at the date of onset of the first non-Raynauds symptom and ending at 15 years, the patients were evaluated at least once a year for these complications. The population used for genotyping of the two PBEF functional variants was composed of 2740 SSc patients and 1913 healthy controls matched by geographical region, age and gender. Six case-control sets were of European ancestry; a Spanish cohort: 228 SSc patients and 265 controls; a Dutch cohort: 203 SSc patients and 277 controls; a German cohort: 312 SSc patients and 247 controls; a British cohort: 269 SSc patients and 39 controls: an Italian cohort 323 SSc patients and 89 controls; and 193 Swedish patients and 167 controls. In addition, we included two distinct ethnic SSc cohorts resident in the USA in the *PBEF* -1001T>G and *PBEF* -1543C>T genotyping; 1047 Caucasian American SSc patients and 692 matched controls and 165 Hispanic American SSc patients and 137 matched controls (**Table 1**).

Table 1. The basic and clinical characteristics of the 8 SSc cohorts included in the genotyping.

Phenotype	Population							
	Dutch	Spanish	German	Swedish	Italian	UK	American white	American Hispanic
n	203	228	312	193	323	269	1047	165
Age (years, SD)	58 (13)	58 (13)	57(12)	53 (14)	55(13)	54 (12)	-	-
Disease duration (months,SD)	131 (82)	144 (90)	113 (109)	82 (73)	140 (138)	155 (92)	-	-
Female %	71.0	83.8	88.4	76.9	94.3	81.2	73.9	74.4
lcSSc %	78.5	70	54.5	82.4	72.3	74.2	60.8	41.1
ATA+ %	22.1	18.4	26.7	16.9	32.9	12.4	16,5	22
ACA+ %	26.2	46.7	41.4	28.3	32.1	37.3	29	7
Pulmonary fibrosis on CT scan %	32.3	30.7	37.2	46.2	32.1	43.2	-	-
FVC (<70% predicted) %	26.1	29.1	18.5	21.1	15.3	30.1	-	-
DLCO (<70% predicted)%	33	45.1	50.2	35.8	67.2	11.5	-	-

Protein Expression

PBEF protein was detected using a human EIA kit (Phoenix pharmaceuticals, Burlingame, CA), according to the manufacturer's instructions. Absorption was measured at 450 nm, and data were analyzed using Revelation version 4.22 software (Dynex Technologies, Denkendorf, Germany). Since PBEF levels have been suggested to be influenced by obesity and gender, only female SSc patients included and the effect of body mass index (BMI) on PBEF expression levels was assessed by regression analysis (14,15).

Genotyping

We assessed two functional polymorphisms in the PBEF promoter region for their role in SSc susceptibility and clinical phenotype. Genotyping was performed using 5' nuclease (Taqman) assays with custom primers and probes (Applied Biosystems, Foster City, CA). For *PBEF* -1001T>G (rs9770242) the forward primer sequence was 5'-ACGGGCCAAGCCTTTGA-3', reverse primer sequence was 5'-CCAACTCGTTTCCCAGGATTTAAAG-3', and reporter sequence was 5'-TCAGTG/TTCGCACCCTG-3'. Corresponding *PBEF* -1543C>T sequences were 5'-GCAAAGATCATGGAAGTGGAAGGTA-3', 5'-CCTCGTTGCTGAAAATAATTGTAAGTGT-3', and 5'-CACCAG/AGCACTCAC-3'.

Data analysis

Significance levels were calculated with 2x2 contingency tables and Fisher's exact test by using SPSS 16.0. Homogeneity of OR among cohorts was calculated using Breslow-Day and Woolf Q methods and the calculation of the pooled OR was performed under a fixed-effects model (Mantel-Haenszel meta-analysis). For the power calculation in the pooled analysis of the SSc patients (n=2737), we considered a minor allele frequency of 0.23 for both variants. This study reaches a power of 80% to detect the effect of a polymorphism at an OR of 1.17 in a log additive model, and a power of 99% at an OR of 1.23. For analysis of quantitative traits the student's t test was applied. Survival analysis was performed using Kaplan-Meier curves and significance levels were calculated with Log Rank (Mantel-Cox) statistics. Cox Proportional Hazards Survival Regression was used to determine relative risks. Combined effects were investigated by multivariate regression analyses. Multiple testing was not applied to the PBEF expression

analyses, since these represent only two separate, subsequent and planned comparisons, namely the validation of previously reported increased PBEF levels in PAH positive SSc patients and validation of previously reported expression influencing polymorphisms. On the results of genotyping of the 8 separate cohorts we applied a Bonferroni correction with a threshold of 0.01, which was based on tests for allele (both polymorphisms), genotype (both polymorphisms) and the haplotype with no minor variants, totaling 0.05/5=0.01. It has to be noted that, this is somewhat too strict since these comparisons are not fully independent. Capped lines represent standard error of the mean throughout the figures in the manuscript. Linkage disequilibrium was assessed using Haploview V4.2 software.

Results

Firstly, we investigated whether PBEF protein levels are indeed higher expressed in SSc patients with PAH, as postulated previously on the mRNA level (2). When we compared the expression levels of PBEF in patients with (n=10) and without (n=30) PAH, we found a significant increase (P=0.02) of PBEF in those patients with PAH (mean 27.41 SEM 1.92) compared to no PAH (mean 20.26 SEM 0.65). Noteworthy, this population only included non-obese, female subjects, since being female and/or obese has been described as a factor influencing PBEF expression (12,13). We did not find any significant effect of having pulmonary fibrosis, a low FVC% or DLCO% predicted on PBEF expression (**Figure 1a**).

Secondly, we hypothesized that two polymorphisms, *PBEF* -1543C>T and *PBEF* -1001T>G, previously associated with pulmonary disease and able to influence PBEF expression, might be involved in the development of PAH as well. We therefore validated whether these polymorphisms really were able to affect expression levels in 39 SSc patients. Indeed, we found a significant effect of the *PBEF* -1001G allele on expression levels (p=0.02). A similar trend was observed for the minor variant of the *PBEF* -1543 variant, although this did not reach statistical significance. When overlooking the results we theorized that there might be an additive effect of the two minor variants on expression of *PBEF*. For this reason we compared PBEF expression between the patients that carried no or ≥ 1 minor allele. The latter analysis revealed that PBEF was highest in those patients expressing no minor variants (p<0.001) (**Figure 1b**). To untangle the effect of the various identified factors with an impact on PBEF expression, we performed a multivariate regression analysis with the variants coming forward from the above mentioned univariate tests. This analysis showed that both PAH (p=0.018 beta=0.35) and carrying minor variants (p=0.007 beta=-0.407 have their own independent significant effect on PBEF expression levels. The model is displayed in **table 2**. However, we did observe a significant correlation between PAH and carrying minor variants (Pearson correlation = -0.327 P=0.23).

Table 2. Multivariate regression model for the effect of pulmonary arterial hypertension and minor variants on PBEF protein expression

	B	Std. Error	Beta	t	*P*
(Constant)	28.332	3.036		9.333	0
PAH	4.324	1.742	0.35	2.482	0.018
minor variants	-4.716	1.633	-0.407	-2.888	0.007

These observations on the genetic and functional level propelled us to further investigate the role of functional variants in *PBEF* with special emphasis on PAH development. When we compared the development of PAH (total events n=40, population at risk is 964) between patients carrying no expression affecting variants (*PBEF* -1543CC and -1001TT genotype) and patients carrying one or two minor variants we found that patients with no variants had a lower frequency of PAH incidence compared to carriers of either the *PBEF* -1001G minor allele or the PBEF-1543T after 15 years. However only carriers of the *PBEF* -1001G were at higher risk for developing PAH (Log Rank P=0.02, HR 2.38, 95%CI: 1.2-5.3) (**Figure 1c**).

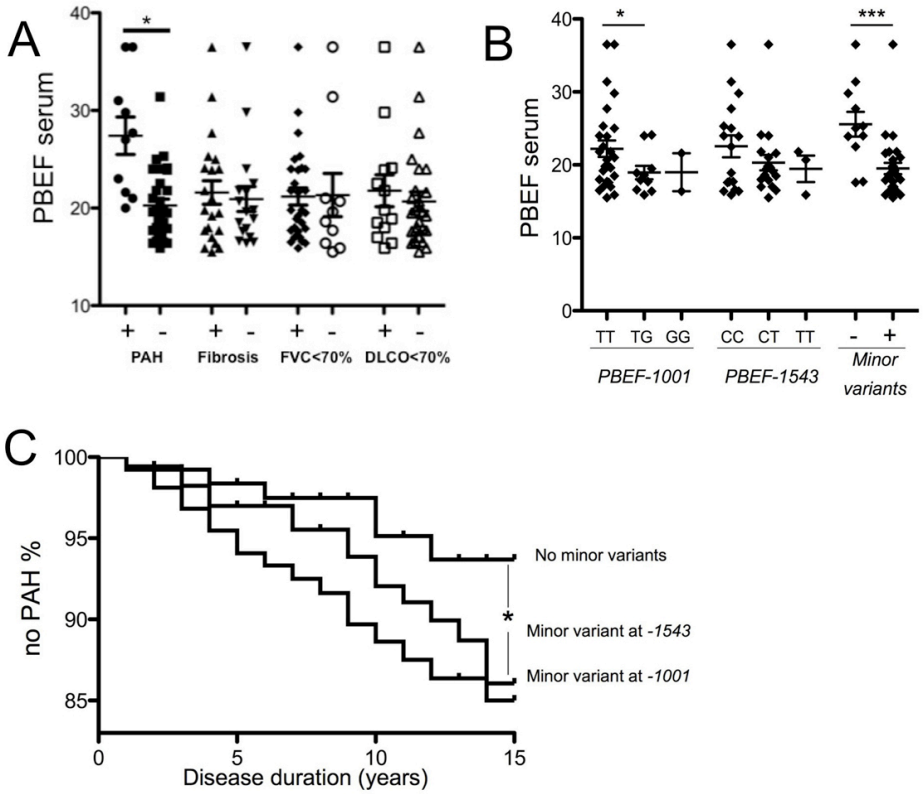


Figure 1. A) PBEF is uniquely higher expressed in SSc patients with PAH (n=10) compared to patients without PAH (n=30) (P=0.02) B) Carriers of minor variants at PBEF-1001 express significantly less PBEF protein in their serum (P=0.02), when carrying no minor variants at PBEF promoter positions -1543 and -1001 PBEF expression was most notably attenuated (P<0.001) C) Kaplan-Meier curve showing a significant increase of PAH development during 15 years follow-up in the patients carrying one or more minor variants at the PBEF-1001 position compared to carriers of no minor alleles (Log Rank P=0.02, HR 2.38, 95%CI: 1.2-5.3). Carrying a minor allele at the PBEF -1543 position did not significantly influence susceptibility for PAH in time, although the total number of PAH cases is higher in this genetic subgroup as well.

Next, we were interested whether these variants have an effect on SSc susceptibility overall. Hardy Weinberg Equilibrium was within proportions in all populations separately. Overall linkage disequilibrium (R2) between the two SNPs is 0.095 and was not significantly different between the populations. Based on the observation that patients with PAH more often carry no minor variants at the *PBEF* -1543 and -1001 promoter sites we were interested whether this combination could also influence SSc susceptibility overall. To analyze this, we compared the frequency of subjects being homozygous for the major alleles of both variants (*PBEF* -1543CC and -1001TT genotype) in SSc patients and healthy controls. This

led to the observation that the lack of both minor alleles markedly increased the risk to develop SSc, in the Swedish (p=0.007 OR 1.94: (95%CI: 1.21-3.10)) and Italian (p=0.02 OR 1.94 (95%CI: 1.10-3.43) population. In a meta-analysis including all populations, we confirmed this association (p =0.009 OR 1.20 (95% CI 1.05-1.37) (**Figure 2**). After applying the Bonferroni correction (treshold p=0.01) this result remains significant. Altogether, variations in the promoter region of *PBEF* seem protective against developing SSc.

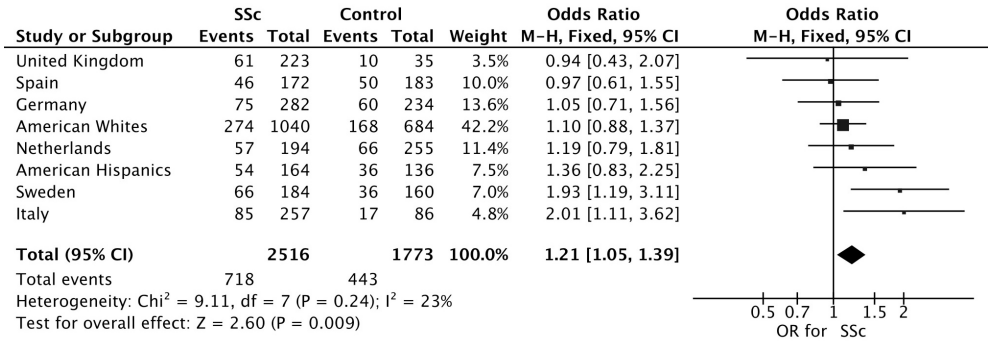


Figure 2 Effects of the combined *PBEF* -1001TT and -1543CC genotypes on SSc susceptibility.

Discussion

In this study we show that PBEF is higher expressed in SSc patients with PAH, the expression is influenced by two polymorphisms in the *PBEF* promoter region, and that these variants together play a role in the risk for developing SSc and PAH.

Although no study investigated the combined effects of *PBEF* -1001 and -1543 variants on PBEF expression, there have been several interesting studies scrutinizing the variants separately. The two *PBEF* promoter variants investigated in this study have previously been associated with lower serum levels of PBEF. More specifically, patients with vascular disease, carrying minor variants at *PBEF* -1543 (annotated as *PBEF* -1535 in the referenced study), had significantly lower levels of PBEF in their plasma similar to the trend observed in our and other studies. Intriguingly, this study also shows that carriers of this minor variant have decreased levels of hs-CRP, interleukin 6 and TNF-α (16). Ineffective up regulation of these molecules due to low PBEF levels might be a partial explanation why carrying minor variants is a protective feature for SSc susceptibility, mainly by affecting its inflammatory component. On the other hand, this variant seemed to attenuate binding of an interleukin 1-beta induced, yet unidentified, transcriptionfactor to PBEF in pulmonary endothelial cells (7). The *PBEF* -1001 minor variant has been shown to reduce PBEF mRNA expression levels in visceral fat and plasma PBEF protein expression, which is in line with our findings as well, although it has to be noted that one study investigating the effect of this polymorphism in acute lung injury patients did not observe a significant effect of this polymorphism on PBEF protein levels (7, 17, 18).

Whereas the combination of the *PBEF* -1543CC and *PBEF* -1001TT genotypes contributes to the risk of developing SSc, this particular combination of genotypes seems to be protective for the development of PAH after onset of SSc. These findings might reflect the pleiotrophic actions of the PBEF molecule. On one hand PBEF is able to influence vascular homeostasis, including permeability of the pulmonary arteries, on the other hand it is known as a strong pro-inflammatory protein (5-7). In regard to this study it translates into the finding that carriers of no variants are able to up regulate PBEF efficiently, these subjects would be more

prone to the pro-inflammatory effects of PBEF but also more able to benefit from its pulmonary regulatory effects. PBEF has not previously been implicated in PAH pathogenesis and is not directly related to the few previously identified risk factors. Nevertheless, our finding is in line with a report describing higher DNA transcription levels of PBEF in leukocytes of SSc patients with PAH (2). However, this hypothesis will need further substantiation in future functional experiments. Noteworthy, these polymorphisms have not been found in a recent genome wide association study in SSc, possibly attributable to the fact that no haplotype analysis was performed with the genome wide data yet and that these polymorphisms are not present as markers on the GWAS genotyping platform used. (13)

Several studies have showed an association of PAH with different genetic variants, but did not correct for disease duration (19, 20). This is a potential bias since the longer disease is present the proportion of patients with PAH will increase as well. In addition, these studies did not only include patients diagnosed with PAH on the basis of right heart catheterization, but also included those with PAH based on isolated DLCO<50% predicted and with unexplained dyspnoea. These criteria are prone to introduce bias, because they are not fully specific for PAH. In this light, this is the first large study that implicates a functional polymorphism in the development of PAH in time using follow up data and robust diagnostic methods for PAH evaluation.

This study harbors some weaknesses as well. First, the relatively small sample size used for the PBEF expression analysis may under- or overestimate the results of this analysis. Second, the expression levels were only investigated in female patients, which are the vast majority of SSc patients. This makes the results not automatically applicable to male SSc patients. Although the results of the meta-analysis survive Bonferroni correction, it seems that the Swedish and Italian cohort mostly produce the observed significant effect. For this purpose a large replication study in an independent cohort is of high interest. In addition, the present study lacks the direct biomechanical evidence necessary to understand the exact role of this protein in SSc and PAH.

Taken together, here we identify the *PBEF* gene as a novel factor that is implicated in the development of SSc related PAH and influences the susceptibility for SSc as well. However, further research is warranted to identify the precise functional and mechanistic processes involved.

References

1. Gabrielli A, Avvedimento EV, Krieg T. Scleroderma. N Engl J Med 2009; 360(19): 1989-2003.

2. Pendergrass SA, Hayes E, Farina G, Lemaire R, Farber HW, Whitfield ML, Lafyatis R. Limited systemic sclerosis patients with pulmonary arterial hypertension show biomarkers of inflammation and vascular injury. PLoS One. 2010 Aug 17;5(8):e12106.

3. Moschen AR, Kaser A, Enrich B, Mosheimer B, Theurl M, Niederegger H et al. Visfatin, an adipocytokine with proinflammatory and immunomodulating properties. J Immunol 2007; 178(3):1748-58.

4. Scala E, Pallotta S, Frezzolini A, Abeni D, Barbieri C, Sampogna F et al. Cytokine and chemokine levels in systemic sclerosis: relationship with cutaneous and internal organ involvement. Clin Exp Immunol 2004; 138(3):540-6.

5. Li H, Liu P, Cepeda J, Fang D, Easley RB, Simon BA et al. Augmentation of Pulmonary Epithelial Cell IL-8 Expression and Permeability by Pre-B-cell Colony Enhancing Factor. J Inflamm (Lond) 2008; 5:15.

6. Liu P, Li H, Cepeda J, Zhang LQ, Cui X, Garcia JG et al. Critical role of PBEF expression in pulmonary cell inflammation and permeability. Cell Biol Int 2009; 33(1): 19-30.

7. Ye SQ, Simon BA, Maloney JP, Zambelli-Weiner A, Gao L, Grant A et al. Pre-B-cell colony-enhancing factor as a potential novel biomarker in acute lung injury. Am J Respir Crit Care Med 2005; 171(4):361-70.

8. Ye SQ, Zhang LQ, Adyshev D, Usatyuk PV, Garcia AN, Lavoie TL et al. Pre-B-cell-colony-enhancing factor is critically involved in thrombin-induced lung endothelial cell barrier dysregulation. Microvasc Res 2005; 70(3):142-51.

9. Bajwa EK, Yu CL, Gong MN, Thompson BT, Christiani DC. Pre-B-cell colony-enhancing factor gene polymorphisms and risk of acute respiratory distress syndrome. Crit Care Med 2007; 35(5):1290-5.

10. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. Arthritis Rheum 1980; 23(5): 581-90.

11. LeRoy EC, Black C, Fleischmajer R, Jablonska S, Krieg T, Medsger TA, Jr. et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. J Rheumatol 1988; 15(2):202-5.

12. Gorlova O, Martin JE, Rueda B et al. Identification of novel genetic markers associated with clinical phenotypes of systemic sclerosis through a genome-wide association strategy. PLoS Genet. 2011 Jul; 7(7):e1002178. Epub 2011 Jul 14.

13. Radstake TR, Gorlova O, Rueda B et al. Genome-wide association study of systemic sclerosis identifies CD247 as a new susceptibility locus. Nat Genet. 2010 May; 42(5):426-9. Epub 2010 Apr 11.

14. Berndt J, Kloting N, Kralisch S, Kovacs P, Fasshauer M, Schon MR et al. Plasma visfatin concentrations and fat depot-specific mRNA expression in humans. Diabetes 2005; 54(10):2911-6.

15. Ons E, Gertler A, Buyse J, Lebihan-Duval E, Bordas A, Goddeeris B et al. Visfatin gene expression in chickens is sex and tissue dependent. Domest Anim Endocrinol 2010; 38(2):63-74.

16. Wang LS, Yan JJ, Tang NP, Zhu J, Wang YS, Wang QM, Tang JJ, Wang MW, Jia EZ, Yang ZJ, Huang J. A polymorphism in the visfatin gene promoter is related to decreased plasma levels of inflammatory markers in patients with coronary artery disease. Mol Biol Rep. 2011 Feb;38(2): 819-25. Epub 2010 Apr 11.

17. Böttcher Y, Teupser D, Enigk B, Berndt J, Klöting N, Schön MR, Thiery J, Blüher M, Stumvoll M, Kovacs P. Genetic variation in the visfatin gene (PBEF1) and its relation to glucose metabolism and fat-depot-specific messenger ribonucleic acid expression in humans. J Clin Endocrinol Metab. 2006 Jul; 91(7):2725-31. Epub 2006 Apr 24

18. Carrero JJ, Witasp A, Stenvinkel P, Qureshi AR, Heimbürger O, Bárány P, Suliman ME, Anderstam B, Lindholm B, Nordfors L, Schalling M, Axelsson J. Visfatin is increased in chronic kidney disease patients with poor appetite and correlates negatively with fasting serum amino acids and

triglyceride levels. Nephrol Dial Transplant. 2010 Mar;25(3):901-6. Epub 2009 Nov 30.

19. Wipff J, Dieudé P, Guedj M, Ruiz B, Riemekasten G, Cracowski JL, Matucci-Cerinic M, Melchers I, Humbert M, Hachulla E, Airo P, Diot E, Hunzelmann N, Caramaschi P, Sibilia J, Valentini G, Tiev K, Girerd B, Mouthon L, Riccieri V, Carpentier PH, Distler J, Amoura Z, Tarner I, Degano B, Avouac J, Meyer O, Kahan A, Boileau C, Allanore Y. Association of a KCNA5 gene polymorphism with systemic sclerosis-associated pulmonary arterial hypertension in the European Caucasian population. Arthritis Rheum. 2010 Oct;62(10):3093-100.

20. Dieudé P, Guedj M, Wipff J, Ruiz B, Riemekasten G, Matucci-Cerinic M, Melchers I, Hachulla E, Airo P, Diot E, Hunzelmann N, Cabane J, Mouthon L, Cracowski JL, Riccieri V, Distler J, Meyer O, Kahan A, Boileau C, Allanore Y. Association of the TNFAIP3 rs5029939 variant with systemic sclerosis in the European Caucasian population. Ann Rheum Dis. 2010 Nov;69(11):1958-64. Epub 2010 May 28.

Supplementary table 1. Genotype and allele frequencies of the *PBEF* 1001T>G polymorphism and Mantel-Haenzel analysis for combined effects

Cohort	Phenotype	n	TT	TG	GG	P vs controls	G	P vs controls
Italy	**SSc**	323	0.58	0.35	0.07	0.78	0.24	0.70
	lcSSc	213	0.58	0.36	0.06	0.85	0.24	0.65
	dcSSc	84	0.66	0.27	0.07	0.41	0.21	0.71
	control	89	0.59	0.36	0.05		0.22	
Sweden	**SSc**	193	0.61	0.36	0.03	0.55	0.21	0.39
	lcSSc	117	0.6	0.33	0.07	0.88	0.23	0.90
	dcSSc	50	0.57	0.43	0	0.19	0.21	0.54
	control	167	0.58	0.36	0.06		0.24	
Spain	**SSc**	228	0.57	0.36	0.07	0.80	0.25	0.48
	lcSSc	154	0.57	0.37	0.06	0.64	0.25	0.46
	dcSSc	58	0.55	0.35	0.1	0.95	0.28	0.89
	control	265	0.55	0.36	0.09		0.27	
Germany	**SSc**	312	0.56	0.36	0.08	0.20	0.27	0.08
	lcSSc	163	0.51	0.37	0.12	0.01	0.31	0.01
	dcSSc	121	0.62	0.33	0.05	0.91	0.21	0.71
	control	247	0.6	0.35	0.05		0.22	
United Kingdom	**SSc**	269	0.55	0.37	0.08	0.74	0.27	0.44
	lcSSc	172	0.54	0.37	0.09	0.81	0.27	0.51
	dcSSc	59	0.57	0.36	0.07	0.65	0.25	0.34
	control	39	0.48	0.41	0.11		0.31	
The Netherlands	**SSc**	201	0.59	0.36	0.05	0.47	0.23	0.94
	lcSSc	129	0.6	0.33	0.07	0.99	0.23	0.92
	dcSSc	49	0.53	0.45	0.02	0.13	0.24	0.73
	control	277	0.61	0.32	0.07		0.23	
Caucasian Americans	**SSc**	1047	0.55	0.38	0.07	0.90	0.26	0.59
	lcSSc	604	0.55	0.38	0.07	0.90	0.26	0.62
	dcSSc	434	0.58	0.35	0.07	0.63	0.25	0.46
	control	690	0.55	0.38	0.07		0.27	
Hispanic Americans	**SSc**	165	0.5	0.43	0.07	0.49	0.29	0.30
	lcSSc	66	0.57	0.35	0.08	0.27	0.25	0.11
	dcSSc	94	0.46	0.47	0.07	0.64	0.31	0.67
	control	136	0.46	0.43	0.11		0.33	
Total (Mantel-Haenszel)	**SSc**	2738	0.56	0.37	0.07	0.79	0.25	0.58
	lcSSc	1618	0.56	0.37	0.08	0.76	0.26	0.94
	dcSSc	949	0.57	0.36	0.06	0.31	0.24	0.21
	control	1910	0.56	0.37	0.07		0.26	

Supplementary table 2. Genotype and allele frequencies of the *PBEF* 1543 C>T polymorphism and Mantel-Haenzel analysis for combined effects

Population	Phenotype	n	CC	CT	TT	P vs controls	T	P vs controls
Italy	**SSc**	321	0.64	0.3	0.06	0.03	0.21	0.09
	lcSSc	213	0.66	0.27	0.07	0.01	0.2	0.07
	dcSSc	82	0.57	0.4	0.03	0.39	0.21	0.23
	control	89	0.5	0.45	0.05		0.27	
Sweden	**SSc**	192	0.53	0.39	0.08	0.15	0.27	0.18
	lcSSc	118	0.57	0.34	0.09	0.20	0.26	0.39
	dcSSc	50	0.54	0.39	0.07	0.32	0.27	0.38
	control	165	0.57	0.4	0.03		0.23	
Spain	**SSc**	228	0.58	0.34	0.08	0.40	0.25	0.28
	lcSSc	154	0.58	0.33	0.09	0.32	0.26	0.24
	dcSSc	58	0.6	0.34	0.06	0.99	0.23	0.78
	control	265	0.61	0.34	0.05		0.22	
Germany	**SSc**	310	0.6	0.35	0.05	0.82	0.23	0.55
	lcSSc	163	0.62	0.33	0.05	0.62	0.21	0.33
	dcSSc	121	0.56	0.38	0.06	0.95	0.25	0.74
	control	247	0.58	0.36	0.06		0.24	
United Kingdom	**SSc**	267	0.62	0.33	0.05	0.33	0.22	0.20
	lcSSc	171	0.63	0.33	0.04	0.33	0.21	0.28
	dcSSc	58	0.6	0.32	0.08	0.32	0.25	0.14
	control	39	0.73	0.22	0.05		0.16	
The Netherlands	**SSc**	203	0.6	0.35	0.05	0.30	0.23	0.12
	lcSSc	130	0.56	0.39	0.05	0.50	0.24	0.33
	dcSSc	50	0.63	0.31	0.06	0.55	0.22	0.29
	control	277	0.53	0.39	0.08		0.27	
Caucasian Americans	**SSc**	1046	0.59	0.36	0.05	0.51	0.22	0.35
	lcSSc	603	0.59	0.36	0.05	0.74	0.23	0.50
	dcSSc	433	0.59	0.37	0.04	0.33	0.22	0.46
	control	690	0.58	0.36	0.06		0.24	
Hispanic Americans	**SSc**	164	0.71	0.27	0.02	0.91	0.16	0.67
	lcSSc	66	0.70	0.28	0.02	0.70	0.16	0.73
	dcSSc	93	0.72	0.25	0.03	0.89	0.16	0.77
	control	137	0.73	0.25	0.02		0.15	
Total (Mantel-Haenszel)	**SSc**	2731	0.60	0.34	0.05	0.52	0.23	0.37
	lcSSc	1618	0.61	0.34	0.06	0.59	0.23	0.50
	dcSSc	945	0.60	0.36	0.05	0.63	0.22	0.50
	control	1909	0.59	0.36	0.06		0.23	

Chapter 9

Identification of a functional epistatic 3-locus model that is associated with Systemic Sclerosis

Beretta L*, Broen JCA*, van Bon L, Cossu M, Rueda B, Simeón CP, Vicente E, Castellvi I, Spanish Systemic Sclerosis group, Coenen MJH, Vonk M, Mayo M, Airo' P, Scorza R, Martin J and Radstake TRDJ

Submitted

Abstract

Objective: To investigate whether epistatic interactions of immunity modulating genes influence the susceptibility to Systemic Sclerosis (SSc) and clinical phenotypes

Methods: Five-hundred-eighty-eight systemic sclerosis patients and 580 ethnically matched healthy controls recruited from 3 cohorts (Italy, Netherlands, Spain) were genotyped for 19 single nucleotide polymorphisms (SNPs) in 9 immunity mediating genes. Single-locus and gene-gene interaction analysis were conducted via the chi-square test and via the multifactor dimensionality reduction method. Functional experiments were then conducted on peripheral blood mononuclear cells (PBMCs) isolated from healthy individuals. Quantitative real-time polymerase-chain-reaction (RT-PCR) was used to asses specific mRNA levels after stimulation with specific agonists, according to the epidemiological results.

Results: None of the studied SNPs was statistically associated with SSc or SSc subsets. A 3-factor epistatic model involving the rs187084 SNP in the *toll-like receptor 9 (TLR9)* gene, the rs1800587 SNP in the *interleukin 1-α (IL-1α)* gene and the rs9770242 SNP in the *pre-B cell-enhancing factor (PBEF)* gene was significantly associated with the occurrence of the diffuse cutaneous subset (dcSSc) ($p<0.05$ after 100-fold permutation testing and Bonferroni adjustment). In functional experiments, PBMCs incubated with 10 ng/mL IL-1α and with 4 μg/mL CpG-DNA (a TLR agonist), but not with any of the single agents alone, markedly over-expressed PBEF mRNA, thus demonstrating the functional interlink among these genes.

Conclusion: The epidemiological as well as the functional results we provide may help to explain the complex genetic architecture of SSc and to unveil its subtle pathogenesis.

Introduction

Systemic sclerosis (SSc) is a complex autoimmune disease characterized by fibrosis of the skin, widespread vasculopathy and involvement of internal organs. The pathogenesis of SSc involves the interplay between inflammation and immune system activation, endothelial dysfunction and overt fibrosis (1). Common polymorphisms in immune regulatory genes have been identified which influence the susceptibility to SSc and its clinical phenotypes consistently throughout different populations (2-9). However, other associations were not validated (10-18) or showed inconsistent results (19-26). The inability to replicate an initial association may stem from different reasons including study design, population stratification or phenotype definition (27, 28). As recently suggested, when main effects fail to replicate, gene-gene interaction (epistasis) should also be considered as a potential source of variance (29). Considerable evidence has been published demonstrating the importance of epistasis in complex human diseases, which is considered to be an ubiquitous and fundamental component of the genetic architecture of common diseases (30).

From a statistical point of view, the study of epistasis is challenging and hindered by several computational limitations, including the problem of the sparseness of data into the multidimensional space (31), the loss of power when adjusting for multiple testing to decrease type I error (32, 33) and the loss of power in presence of multicollinearity (34). To circumvent these issues several methods have been developed and/or applied to detect gene-gene or gene-environment interactions in genetic-association studies (http://www.epistasis-list.org). Each of these computational approaches has its own strengths and weaknesses, however, regardless of the method, it is well recognized that it is difficult to make biological inferences from statistical models of epistasis (34). Indeed, high- and low-risk assignments that result either from generative (e.g. conventional statistic) or discriminative (e.g. data-mining) modeling usually do not show a clear trend and distribution across the multidimensional space and are not-easily interpretable. A solution to this problem has been proposed by Moore and co-workers (35); here, information theory measures were used to gauge the relative contribution of epistatic attributes to statistical models of gene-gene interaction and graphical models (36) were used to quickly identify additive and non-additive interaction

effects. Nonetheless, to which extent these graphical and mathematical models would fit the biological process of epistasis at the cellular or organism level is presently unknown. As a consequence, high-order epistatic models described in the literature so far, do merely report an association between high-risk multilocus combinations and diseases without validating their findings in *in vitro* or *in vivo*. In the present work we show that in a candidate-gene case-control association study conducted in a large population of SSc patients, gene-gene interaction may be more informative than single-locus analysis, providing, at one time functional evidence for the epistatic model we describe.

Methods

Patients and Controls

Five-hundred-fifty-eight SSc patients recruited from 3 cohorts (Italy: 207; Netherlands: 145; Spain: 236 patients) and 580 healthy ethnically and geographically-matched controls were considered. All the patients fulfilled the American College of Rheumatology 1980 classification criteria for the classification of SSc (37) and were classified as having the limited cutaneous (lcSSc) or the diffuse cutaneous (dcSSc) subset, according to LeRoy et al (38) and the patients' autoantibody profile was determined by reviewing the patients' medical records. Disease onset was defined on the basis of the patient's recall of the first non-Raynaud feature clearly attributable to SSc (39)

The study was approved by the local ethics committees and both patients and controls provided written informed consent for the research.

Genotyping

Nineteen single-nucleotide polymorphisms (SNPs) in 9 immunomodulatory genes were genotyped both in patients and controls via the TaqMan SNP genotyping assays (Applied Biosystems): *toll-like receptor (TLR) 2*, rs1898830 and rs5743704; *TLR4*, rs4986790, rs4986791 and, rs7873784; *TLR9*, rs187084; p*re-B cell-enhancing factor (PBEF)*, rs9770242 and -1543 C>T; *interleukin (IL) 1α*, rs2856838, rs1800587, rs3783557 and rs3783525; CD69, rs17806015, rs6416258, rs11052877 and rs12368445; *adenosine monophosphate deaminase-1 (AMPD1)* rs17602729; *IL4* rs2243250; *IL13* rs20541. Primers for the studied SNPs are available upon request to the authors.

Statistical analysis

The distribution of genotypes was tested for Hardy-Weinberg equilibrium with the goodness-of-fit χ^2 test. Missing genotypes were imputed via the k-nearest neighbor method (40) using the Orange data-mining suite (available at: http://www.ailab.si/orange/), whenever less than 20% of genotypes was available, subjects were excluded. Single SNP association analysis performed using the χ^2 test or the Fischer's exact test (when the expected cell count was below 5); results were

declared significant at the 0.05 threshold after Bonferroni adjustment fro multiple comparisons. To assess the combined effect of the studied SNPs in all populations, we also performed a meta-analysis using Mantel-Haenszel statistics with fixed effects, checking that the percentage of total variation across studies that is due to heterogeneity (I^2) was below 50% (41). Gene-gene interaction analysis was performed using the multifactor dimensionality reduction (MDR) algorithm (42, 43) (available at: http://www.epistasis.org). The MDR algorithm was optimized with a "scaling" approach (sMDR) incorporating LD as described in the Supplementary Notes. Three-thousand synthetic datasets and 24000 runs of the algorithm were used to determine the specific sMDR power (after 100-fold permutation testing and Bonferroni adjustment). For gene-gene interaction analysis all the three available populations were pooled together to increase the power to detect epistasis (44). Before pooling, we verified that the marginal distribution of the genetic variants did not differ across ethnicities, which makes biases unlikely (45). We further discarded the existence of stratification by applying the methodology described by Pritchard and Rosenberg (46); in this case, only the candidate loci themselves were used in the test for stratification. Finally, when more SNPs were in linkage disequilibrium (LD), the SNP/LD block with the highest χ^2 test value, that is the SNP that contributed most to differences across ethnicities, was chosen for the stratification tests.

Functional study, cell cultures.

To provide an interpretation of gene-gene interactions models we used measures of interaction theory and interactions graphs to visualize the nature of the dependencies of the attributes included in the best epistatic model. All these features are implemented in the MDR package. Graphical models were then interpreted in relation to the current knowledge about the involved genes and functional experiments conducted accordingly (see results). Experiments were conducted on peripheral blood mononuclear cells (PBMCs) isolated by Ficoll-Hystopaque density gradient centrifugation from 6 healthy individuals. Cells were incubated for 12 hours with 10 ng/mL IL-1α for 24 hours or with 4 μg/mL CpG-DNA, a specific TLR9 agonists (47). In parallel, PBMCs were pre-stimulated with 10 ng/ml IL-1α for 12 hours after which PBMCs were co-cultured with 4 μg/mL CpG-DNA for another 12 hrs. In all these experiments PBEF mRNA expression was

measured using real-time polymerase-chain reaction (RT-PCR). Delta CT between GAPDH and PBEF of the stimulated cells were compared with the dCT of the unstimulated cells. All the experiments were conducted in triplicate.

Results

Clincal and demographic characteristics of the three SSc cohorts are reported in **Table 1.** After genotyping, imputation and exclusion of instances with low call rate, the dataset was reduced to 437 SSc patients and 506 controls. In this set, single SNP association analysis revealed no significant association with SSc or SSc subsets in the three separate populations. Similarly, meta-analysis revealed no significant association with any of the studied SNP.

Table 1. Demographics

Variable	Netherlands (n=145)	Spain (n=236)	Italy (n=207)
Age (years, SD)	54 (12)	55 (14)	53(11)
Disease duration (months,SD)	131 (85)	141 (92)	148 (132)
Females (%)	72.3	87.6	90.0
lcSSc (%)	75.2	72.6	73.4
Positivity anti-topo (%)	23.9	21,8	43.8
Positivity ACA (%)	26.1	55.5	36.3

In absence of detectable main genetic effects, we performed gene-gene interaction analysis. In the pooled population, gene-gene interaction analysis yielded no statistically significant results when comparing SSc with controls or lcSSc with controls. On the contrary, a 3-factor model involving the *TLR9* rs187084, *IL-1α* rs1800587 and the *PBEF* rs9770242 SNP resulted to be predictive of dcSSc at a significance level of 0.025 after 100-fold permutation testing. The epistatic model involving these SNPs had a sensitivity = 0.584 and a specificity = 0.609 to predict dcSSc occurrence in our sample population.

To better gauge the contribution of the abovementioned epistatic attributes, as well as the magnitude of their interacting effects on the dcSSc status, graphical representations based on information theory, as implemented in the MDR package, were performed. There, we observed that, the rs9770242 (*PBEF*) SNP had the highest degree of individual effect on dcSSc susceptibility, however the highest amount of information (e.g. the highest synergy) is obtained joining the rs187084 (*TLR9*) to the rs1800587 (*IL-1α*), genetic variant (**Figure 1**). When indeed *TLR9* and *IL-1α* are interleaved together (35) they provide redundant information with

that provided by the *PBEF* gene. Accordingly, we postulated that *PBEF* would be the main contributing factor to dcSSc occurrence and that IL-1α would modify/ modulate TLR9 signaling to produce PBEF responses.

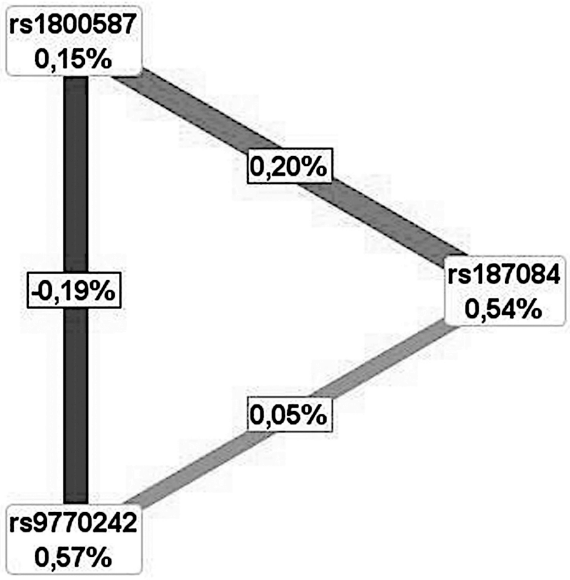


Figure 1 - Interaction graph for the best epistatic model. The interaction model describes the percent of the entropy that is explained by each factor or two-way interaction. The percentage in the node expresses the amount of label's uncertainty eliminated by the node's attribute and the connection the relative mutual information, a red or orange line suggests a positive information gain which can be interpreted as a synergistic or nonadditive relationship while a blue line suggests a loss of information which can be interpreted as redundancy or correlation. The maximum information gain is observed between the rs187084 (toll-like receptor 9) and the rs1800587 (interleukin-1α) single nucleotide polymorphisms.

Functional experiments were driven according to the abovementioned model and, namely, to test the biological effect of *TLR9/IL-1α* interaction on *PBEF*. Here, we observed that the incubation with IL-1α or CpG-DNA alone did not lead to a significantly increased PBEF expression in PBMCs from healthy individuals. In contrast, pre-stimulation with IL-1α and consecutive CpG-DNA-mediated stimulation led to a significant increase of PBEF expression compared to PBMCs that were either not stimulated or stimulated with IL-1α or CpG-DNA alone (p< 0.03, **Figure 2**).

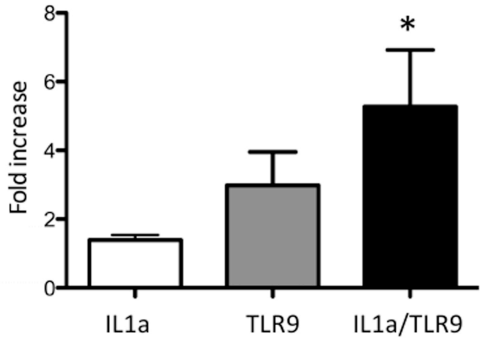


Figure 2 - Interaction effect between IL-1α and TLR9 in PBMCs. Fold-increase of pro-B cell-enhancing factor (PBEF) mRNA expression upon stimulation with either intereukin-1 alpha (IL-1α), toll-like receptor 9 (TLR9) agonist CpG-DNA or prestimulation with IL-1α and subsequent TLR9 stimulation in peripheral blood mononuclear cells (PBMCs). Comparing dCT values, we observed a significant increase in PBEF mRNA production by PBMCs upon a prestimulation with IL-1α and subsequent TLR9 stimulation, compared to the unstimulated PBMCs and PBMCs stimulated with IL-1α alone (p=0.03) in 8 healthy controls.

Discussion

Susceptibility to SSc as well as to other complex human diseases does result form the interplay between several genetic and environmental factors. Whilst a number of single-locus case-controls studies (2-9) as well as a recent genome-wide association study from our collaborative group (48) have been very successful in identifying a handful of specific genetic variants that underly the susceptibility to SSc, it is disappointing that still only a small proportion of its heritability has been accounted for. It is likely that the 'missing heritability' of complex human diseases may be in part explained by epistasis (30, 49, 50). Evidence for such interactions has emerged, for instance, from reworking data from the Wellcome Trust Case–Control Consortium data sets (51) but are still lacking in other autoimmune disease, such as SSc.

The present study is the first successful attempt to fully demonstrate both epidemiologically and functionally, the existence of a genetic biological network in SSc and in multifactorial human diseases. The network we describe is "bi-compartmental", in the sense that we could sort out one main factor (*PBEF)* that interacts with a two-gene complex (*IL-1α/TLR9*). With stimulation experiments in PBMCs we indeed confirmed that co-stimulation with IL-1α and TLR9 agonist is capable of modulating PBEF mRNA expression. The precise mechanism of the IL-1α/TLR9 interaction is presently unknown. Since CpG-DNA alone is not sufficient to boost IL-1α gene expression (52), we speculate that this interaction acts downstream of TLR/IL-1 receptors following specific ligand activation. Whether the net result of this interaction is a qualitative or a quantitative modification of the interleukin-1 receptor associated kinase (IRAK)-pathway and/or of the MyD88 signaling that characterize TLR9 responses (53), needs to be verified in further experiments.

Several lines of evidence suggest that PBEF may be relevant in SSc susceptibility and phenotypical expression. PBEF has a longevity promoting effect on smooth muscle cells, fibroblasts and neutrophils (54). Next to this, PBEF in combination with SIRT1 is able to up regulate fibroblast growth factor 2 (FGF2) (55), a mediator that is found to be abundantly expressed in SSc skin and, together with TGF-β, strongly promotes fibrosis (56). The production of IL6, IL10 and TNFα, that are

found to be increased in SSc (57), can be mounted through PBEF as well (58). Additionally, PBEF may contribute to the deregulation of metalloproteinases (59), a well-known phenomenon that contributes to skin fibrosis in SSc, especially in those patients with the most extensive cutaneous involvement (60). Finally, PBEF may consistently alter endothelial function and vascular homeostasis (61), two pathophysiological processes that play a relevant role in the early stages of fibrosis development (1). Taken together, an enhancement of PBEF expression through the TLR9 and IL-1α interaction could impel the inflammation, matrix remodelling and subsequent fibrosis as observed in dcSSc.

In summary, in the present work we identified a unique epistatic three locus interaction model comprising *IL-1α*, *TLR9* and *PBEF* that is associated with an increased susceptibility to dcSSc. We then constructed a theoretical model to biologically explain the epidemiological correlate of this interaction and verified the model in vitro. Accordingly, we described a novel interaction between IL-1α and TLR9 that promotes *PBEF* gene expression. To our knowledge, this is the first time that a statistical model of gene-gene interaction is shown to be biologically functional in any human disease. The results we provide may help to unveil the complex genetic architecture and pathogenesis of SSc, and highlight new biological pathways playing a role in the pathogenesis of SSc.

References

1. Varga J, Abraham D. Systemic sclerosis: a prototypic multisystem fibrotic disorder. J. Clin. Invest. 2007;117:557–567.

2. Broen J, Gourh P, Rueda B, Coenen M, Mayes M, Martin J, Arnett FC, Radstake TR; European Consortium on Systemic Sclerosis Genetics. The FAS -670A>G polymorphism influences susceptibility to systemic sclerosis phenotypes. Arthritis Rheum. 2009;60:3815-20.

3. Dieudé P, Wipff J, Guedj M, Ruiz B, Melchers I, Hachulla E, Riemekasten G, Diot E, Hunzelmann N, Sibilia J, Tiev K, Mouthon L, Cracowski JL, Carpentier PH, Distler J, Amoura Z, Tarner I, Avouac J, Meyer O, Kahan A, Boileau C, Allanore Y. BANK1 is a genetic risk factor for diffuse cutaneous systemic sclerosis and has additive effects with IRF5 and STAT4. Arthritis Rheum. 2009;60:3447-54.

4. Dieudé P, Guedj M, Wipff J, Ruiz B, Hachulla E, Diot E, Granel B, Sibilia J, Tiev K, Mouthon L, Cracowski JL, Carpentier PH, Amoura Z, Fajardy I, Avouac J, Meyer O, Kahan A, Boileau C, Allanore Y. STAT4 is a genetic risk factor for systemic sclerosis having additive effects with IRF5 on disease susceptibility and related pulmonary fibrosis. Arthritis Rheum. 2009;60:2472-9.

5. Gourh P, Agarwal SK, Divecha D, Assassi S, Paz G, Arora-Singh RK, Reveille JD, Shete S, Mayes MD, Arnett FC, Tan FK. Polymorphisms in TBX21 and STAT4 increase the risk of systemic sclerosis: Evidence of possible gene-gene interaction and alterations in Th1/Th2 cytokines. Arthritis Rheum. 2009;60:3794-3806.

6. Liakouli V, Manetti M, Pacini A, Tolusso B, Fatini C, Toscano A, Cipriani P, Guiducci S, Bazzichi L, Codullo V, Ruocco L, Dell'orso L, Carubbi F, Marrelli A, Abbate R, Bombardieri S, Ferraccioli G, Montecucco C, Valentini G, Matucci-Cerinic M, Ibba-Manneschi L, Giacomelli R. The -670G>A polymorphism in the FAS gene promoter region influences the susceptibility to systemic sclerosis. Ann Rheum Dis. 2009;68:584-90.

7. Rueda B, Broen J, Simeon C, Hesselstrand R, Diaz B, Suárez H, Ortego-Centeno N, Riemekasten G, Fonollosa V, Vonk MC, van den Hoogen FH, Sanchez-Román J, Aguirre-Zamorano MA, García-Portales R, Pros A, Camps MT, Gonzalez-Gay MA, Coenen MJ, Airo P, Beretta L, Scorza R, van Laar J, Gonzalez-Escribano MF, Nelson JL, Radstake TR, Martin J. The STAT4 gene influences the genetic predisposition to systemic sclerosis phenotype. Hum Mol Genet. 2009;18:2071-7.

8. Rueda B, Gourh P, Broen J, Agarwal SK, Simeon C, Ortego-Centeno N, Vonk MC, Coenen M, Riemekasten G, Hunzelmann N, Hesselstrand R, Tan FK, Reveille JD, Assassi S, Garcia-Hernandez FJ, Carreira P, Camps M, Fernandez-Nebro A, Garcia de la Peña P, Nearney T, Hilda D, Gónzalez-Gay MA, Airo P, Beretta L, Scorza R, Radstake TR, Mayes MD, Arnett FC, Martin J. BANK1 functional variants are associated with susceptibility to diffuse systemic sclerosis in Caucasians. Ann Rheum Dis. 2010;69:700-5.

9. Tsuchiya N, Kawasaki A, Hasegawa M, Fujimoto M, Takehara K, Kawaguchi Y, Kawamoto M, Hara M, Sato S. Association of STAT4 polymorphism with systemic sclerosis in a Japanese population. Ann Rheum Dis. 2009;68:1375-6.

10. Beretta L, Santaniello A, Cappiello F, Barili M, Scorza R. No evidence for a role of the proximal IL-6 G/C -174 single nucleotide polymorphism in Italian patients with systemic sclerosis. J Cell Mol Med. 2007;11:896-8.

11. Beretta L, Cossu M, Marchini M, Cappiello F, Artoni A, Motta G, Scorza R. A polymorphism in the human serotonin 5-HT2A receptor gene may protect against systemic sclerosis by reducing platelet aggregation. Arthritis Res Ther. 2008;10:R103.

12. Carulli MT, Spagnolo P, Fonseca C, Welsh KI, duBois RM, Black CM, Denton CP. Single-nucleotide polymorphisms in CCL2 gene are not

associated with susceptibility to systemic sclerosis. J Rheumatol. 2008;35:839-44.

13. Kirsten H, Burkhardt J, Hantmann H, Hunzelmann N, Vaith P, Ahnert P, Melchers I. 5HT2A polymorphism His452Tyr in a German Caucasian systemic sclerosis population. Arthritis Res Ther. 2009;11:403.

14. Lagan AL, Pantelidis P, Renzoni EA, Fonseca C, Beirne P, Taegtmeyer AB, Denton CP, Black CM, Wells AU, du Bois RM, Welsh KI. Single-nucleotide polymorphisms in the SPARC gene are not associated with susceptibility to scleroderma. Rheumatology (Oxford). 2005;44:197-201.

15. Radstake TR, Vonk MC, Dekkers M, Schijvenaars MM, Treppichio WL, Lafyatis R, Riemekasten G, van den Hoogen F, Coenen MJ. The -2518A>G promoter polymorphism in the CCL2 gene is not associated with systemic sclerosis susceptibility or phenotype: results from a multicenter study of European Caucasian patients. Hum Immunol. 2209;70:130-3.

16. Rueda B, Broen J, Torres O, Simeon C, Ortego-Centeno N, Schijvenaars MM, Vonk MC, Fonollosa V, van den Hoogen FH, Coenen MJ, Sanchez-Román J, Aguirre-Zamorano MA, García-Portales R, Pros A, Camps MT, Gonzalez-Gay MA, Martin J, Radstake TR.. The interleukin 23 receptor gene does not confer risk to systemic sclerosis and is not associated with systemic sclerosis disease phenotype. Ann Rheum Dis. 2009;68:253-6.

17. Sfrent-Cornateanu R, Mihai C, Balan S, Ionescu R, Moldoveanu E. The IL-6 promoter polymorphism is associated with disease activity and disability in systemic sclerosis. J Cell Mol Med. 2006;10:955-9.

18. Zhou X, Tan FK, Reveille JD, Wallis D, Milewicz DM, Ahn C, Wang A, Arnett FC. Association of novel polymorphisms with the expression of SPARC in normal fibroblasts and with susceptibility to scleroderma. Arthritis Rheum. 2002;46:2990-9.

19. Beretta L, Bertolotti F, Cappiello F, Barili M, Masciocchi M, Toussoun K, Caronni M, Scorza R. Interleukin-1 gene complex polymorphisms in systemic sclerosis patients with severe restrictive lung physiology. Hum Immunol. 2007;68:603-9.

20. Guiducci S, Fatini C, Georgountzos A, Sticchi E, Cinelli M, Kaloudi O, Rogai V, Melchiorre D, Pignone A, Vlachoyannopoulos P, Abbate R, Matucci Cerinic M. Etrurians vs Greeks: May ACE I/D polymorphism still be considered as a marker of susceptibility to SSc? Clin Exp Rheumatol. 2006;24:432-4.

21. Fonseca C, Lindahl GE, Ponticos M, Sestini P, Renzoni EA, Holmes AM, Spagnolo P, Pantelidis P, Leoni P, McHugh N, Stock CJ, Shi-Wen X, Denton CP, Black CM, Welsh KI, du Bois RM, Abraham DJ. A polymorphism in the CTGF promoter region associated with systemic sclerosis. N Engl J Med. 2007;357:1210-20.

22. Hutyrová B, Lukác J, Bosák V, Buc M, du Bois R, Petrek M. Interleukin 1alpha single-nucleotide polymorphism associated with systemic sclerosis. J Rheumatol. 2004;31:81-4.

23. Kawaguchi Y, Tochimoto A, Ichikawa N, Harigai M, Hara M, Kotake S, Kitamura Y, Kamatani N. Association of IL1A gene polymorphisms with susceptibility to and severity of systemic sclerosis in the Japanese population. Arthritis Rheum. 2003;48:186-92.

24. Kawaguchi Y, Ota Y, Kawamoto M, Ito I, Tsuchiya N, Sugiura T, Katsumata Y, Soejima M, Sato S, Hasegawa M, Fujimoto M, Takehara K, Kuwana M, Yamanaka H, Hara M. Association study of a polymorphism of the CTGF gene and susceptibility to systemic sclerosis in the Japanese population. Ann Rheum Dis. 2009;68:1921-4.

25. Rueda B, Simeon C, Hesselstrand R, Herrick A, Worthington J, Ortego-Centeno N, Riemekasten G, Fonollosa V, Vonk MC, van den Hoogen FH, Sanchez-Román J, Aguirre-Zamorano MA, García-Portales R, Pros A, Camps MT, Gonzalez-Gay MA, Gonzalez-Escribano MF, Coenen MJ, Lambert N, Nelson JL, Radstake TR, Martin J. A large multicentre analysis of CTGF -945 promoter polymorphism does not

confirm association with systemic sclerosis susceptibility or phenotype. Ann Rheum Dis. 2009;68:1618-20.

26. Wipff J, Gallier G, Dieude P, Avouac J, Tiev K, Hachulla E, Granel B, Diot E, Sibilia J, Mouthon L, Meyer O, Kahan A, Varret M, Boileau C, Allanore Y. Angiotensin-converting enzyme gene does not contribute to genetic susceptibility to systemic sclerosis in European Caucasians. J Rheumatol. 2009;36:337-40.

27. Sillanpää MJ, Auranen K. Replication in genetic studies of complex traits. Ann Hum Genet. 2004;68:646-57.

28. NCI-NHGRI Working Group on Replication in Association Studies, Chanock SJ, Manolio T, Boehnke M, Boerwinkle E, Hunter DJ, Thomas G, Hirschhorn JN, Abecasis G, Altshuler D, Bailey-Wilson JE, Brooks LD, Cardon LR, Daly M, Donnelly P, Fraumeni JF Jr, Freimer NB, Gerhard DS, Gunter C, Guttmacher AE, Guyer MS, Harris EL, Hoh J, Hoover R, Kong CA, Merikangas KR, Morton CC, Palmer LJ, Phimister EG, Rice JP, Roberts J, Rotimi C, Tucker MA, Vogan KJ, Wacholder S, Wijsman EM, Winn DM, Collins FS. Replicating genotype-phenotype associations. Nature. 2007;447;655-60.

29. Greene CS, Penrod NM, Williams SM, Moore JH. Failure to replicate a genetic association may provide important clues about genetic architecture. PLoS One. 2009;4:e5639.

30. Moore JH. The ubiquitous nature of epistasis in determining susceptibility to common human diseases. Hum Hered. 2003;56:73–82.

31. Bellman R. Adaptive Control Processes Princeton NJ: Princeton University Press, 1961.

32. Concato J, Feinstein AR, Holford TR. The risk of determining risk with multivariable models. Ann Int Med. 1996;118:201–210.

33. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. Controlling the false discovery rate in behavior genetics research. Behav Brain Res. 2001;125, 279-284.

34. Moore JH, Williams SM. Traversing the conceptual divide between biological and statistical epistasis: systems biology and a more modern synthesis. BioEssays 2005;27:637-646.

35. Moore JH, Gilbert JC, Tsai CT, Chiang FT, Holden T, Barney N, White BC. A flexible computational framework for detecting, characterizing, and interpreting statistical patterns of epistasis in genetic studies of human disease susceptibility. J Theor Biol. 2006;241:252-61.

36. Jakulin A, Bratko I. Analyzing attribute interactions. Lect Notes Artif Intell. 2003;2838:229–40.

37. Subcommittee for Scleroderma Criteria of the American Rheumatism Association. (1980) Diagnostic and Therapeutics Criteria Committee. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Arthritis Rheum 23:581-90.

38. LeRoy EC, Black C, Fleischmajer R, Jablonska S, Krieg T, Medsger TA Jr, Rowell N, Wollheim F. Scleroderma (systemic sclerosis): classification, subset and pathogenesis. J Rheumatol 15:202-5, 1988.

39. White B, Bauer EA, Goldsmith LA, Hochberg MC, Katz LM, Korn JH. Guidelines for clinical trials in systemic sclerosis (scleroderma). I. Disease-modifying interventions. The American College of Rheumatlogy Committee on Design and Outcomes in Clinical Trials in Systemic Sclerosis. Sem Arthritis Rheum 38:351-60, 1995.

40. Troyanskaya O, Cantor M, Sherlock G, Brown P, Hastie T, Tibshirani R, Botstein D, Altman RB. Missing value estimation methods for DNA microarrays. Bioinformatics, 2001;17:520–525.

41. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21:1539-58.

42. Ritchie MD, Hahn LW, Roodi N, Bailey LR, Dupont WD, Parl FF, Moore JH. Multifactor-dimensionality reduction reveals high-order interactions among

estrogen-metabolism genes in sporadic breast cancer. Am J Hum Genet. 2001;69:138-47.

43. Hahn LW, Ritchie MD, Moore JH. Multifactor dimensionality reduction software for detecting gene-gene and gene-environment interactions. Bioinformatics. 2003;19,376-82.

44. Hirschhorn, J.N., Lohmueller, K., Byrne, E., Hirschhorn, K.A. (2002). A comprehensive review of genetic association studies. Genet. Med. 4:45–61.

45. Wang Y, Localio R, Rebbeck TR. Evaluating Bias due to Population Stratification in Epidemiologic Studies of Gene-Gene or Gene-Environment Interactions. Cancer Epidemiol Biomarkers Prev. 2006;15:124-32.

46. Pritchard JK, Rosenberg NA. Use of Unlinked Genetic Markers to Detect Population Stratification in Association Studies. Am J Hum Genet. 1999;65:220–228

47. Takeshita F, Leifer CA, Gursel I, Ishi, KJ, Takeshita S, Gursel M, Klinman DM. Cutting edge: Role of Toll-like receptor 9 in CpG DNA-induced activation of human cells. J Immunol. 2001;167:3555-8.

48. Radstake TR, Gorlova O, Rueda B, Martin JE, Alizadeh BZ, Palomino-Morales R, Coenen MJ, Vonk MC, Voskuyl AE, Schuerwegh AJ, Broen JC, van Riel PL, van 't Slot R, Italiaander A, Ophoff RA, Riemekasten G, Hunzelmann N, Simeon CP, Ortego-Centeno N, González-Gay MA, González-Escribano MF; Spanish Scleroderma Group, Airo P, van Laar J, Herrick A, Worthington J, Hesselstrand R, Smith V, de Keyser F, Houssiau F, Chee MM, Madhok R, Shiels P, Westhovens R, Kreuter A, Kiener H, de Baere E, Witte T, Padykov L, Klareskog L, Beretta L, Scorza R, Lie BA, Hoffmann-Vold AM, Carreira P, Varga J, Hinchcliff M, Gregersen PK, Lee AT, Ying J, Han Y, Weng SF, Amos CI, Wigley FM, Hummers L, Nelson JL, Agarwal SK, Assassi S, Gourh P, Tan FK, Koeleman BP, Arnett FC, Martin J, Mayes MD. Genome-wide association study of systemic sclerosis identifies CD247 as a new susceptibility locus. Nat Genet. 2010 May;42(5):426-9.

49. Clarke AJ, Cooper DN. GWAS: heritability missing in action? Eur J Hum Genet. 2010 Mar 17.

50. Greenspan RJ. The flexible genome. Nat Rev Genet. 2001;2:383-7.

51. Emily M, Mailund T, Hein J, Schauser L, Schierup MH: Using biological networks to search for interacting loci in genome-wide association studies. Eur J Hum Genet 2009; 17: 1231–1240.

52. Gao JJ, Diesl V, Wittmann T, Morrison DC, Ryan JL, Vogel SN, Follettie MT. Bacterial LPS and CpG DNA differentially induce gene expression profiles in mouse macrophages. J Endotoxin Res. 2003;9:237-43.

53. Kim TW, Staschke K, Bulek K, Yao J, Peters K, Oh KH, Vandenburg Y, Xiao H, Qian W, Hamilton T, Min B, Sen G, Gilmour R, Li X. A critical role for IRAK4 kinase activity in Toll-like receptor-mediated innate immunity. J Exp Med. 2007;14;204:1025-36.

54. van der Veer E, Ho C, O'Neil C, Barbosa N, Scott R, Cregan SP, Pickering JG. Extension of human cell lifespan by nicotinamide phosphoribosyltransferase. J Biol Chem. 2007;282:10841-5.

55. Bae YH, Bae MK, Kim SR, Lee JH, Wee HJ, Bae SK. Upregulation of fibroblast growth factor-2 by visfatin that promotes endothelial angiogenesis. Biochem Biophys Res Commun. 2009;379:206-11.

56. Lawrence A, Khanna D, Misra R, Aggarwal A. Increased expression of basic fibroblast growth factor in skin of patients with systemic sclerosis. Dermatol Online J. 2006;27;12:2.

57. Scala E, Pallotta S, Frezzolini A, Abeni D, Barbieri C, Sampogna F, De Pità O, Puddu P, Paganelli R, Russo G. Cytokine and chemokine levels in systemic sclerosis: relationship with cutaneous and internal organ involvement. Clin Exp Immunol. 2004;138:540-6.

58. Moschen AR, Kaser A, Enrich B, Mosheimer B, Theurl M, Niederegger H, Tilg H. Visfatin, an adipocytokine

with proinflammatory and immunomodulating properties. J Immunol. 2007;178:1748-58.

59. Adya R, Tan BK, Chen J, Randeva HS. Nuclear factor-kappaB induction by visfatin in human vascular endothelial cells: its role in MMP-2/9 production and activation. Diabetes Care. 2008;31:758-60.

60. Jinnin M. Mechanisms of skin fibrosis in systemic sclerosis. J Dermatol. 2010;37:11-25.

61. Ye, S.Q., Zhang, L.Q., Adyshev, D., Usatyuk, P.V., Garcia, A.N., Lavoie, T.L., Verin, A.D., Natarajan, V., Garcia, J.G. (2005). Pre-B-cell-colony-enhancing factor is critically involved in thrombin-induced lung endothelial cell barrier dysregulation. Microvasc Res. 70:142-51.

Supplementary notes 1

A scaling approach to multifactor dimensionality reduction (MDR) analysis and power analysis.

The multifactor dimensionality algorithm is by far the most popular method used to detect epistasis in candidate-genes case-controls studies (http://www.epistasis-list.org). In the version proposed by the developers, k-fold cross-validation (where k is usually equal to 5 or 10) is used to sort out the causative genetic variants that epistatically interact within a given dataset (Ritchie et al, 2001). The causative variants are those that maximize the balanced accuracy (BA) function (e.g. the mathematical mean of sensitivity and specificity) in the k training sets. To properly select the interacting model, to reduce overfitting and to estimate its capability of generalization, BA values in the k testing sets are then calculated. The statistical significance of the testing BA values can eventually be computed by the use of resampling techniques, such as permutation testing (Moore et al, 2006; Greene et al, 2009). The combination of cross-validation and permutation testing ensures a reasonable power of the method and minimizes type I errors.

Running MDR in any dataset and tracking the top list of models in a given dimension (a function embedded in the last available version of the software), it is clear that models that do not rank first may have higher testing BA values than the topmost model. Similarly, in synthetic datasets, it may happen that the causative loci are not those with the highest mean training BA values. The inclusion of these model in the analysis could thus theoretically improve the discriminative ability of MDR. Nonetheless, if these model were considered by the user, the background noise would be amplified and the threshold for significance should be corrected for the higher number of tests and become more conservative. The price paid to maintain stable the type I error would be a loss of power. Nevertheless, our working hypothesis is that the inclusion of additional models into the analysis would help to wade through the noise and dampen the negative effect of the increasing number of tests.

To proof our point we developed a modified version of the MDR algorithm, henceforth labelled as scaling MDR (sMDR). In the sMDR method, the thresholds for significance are established before the MDR algorithm is run. This can be accomplished either randomly permuting n-times the studied dataset or creating a

null population with the same characteristics of the latter and then extracting n samples from that. MDR is then run in each of the null samples and h top models ($Top_{1,2,...,h}$) per dimension are sorted out. Testing BA values for the Top_h ranking models are then computed. The testing BA value of any model is corrected whenever the testing BA values of the models with a lower index are higher. Let's suppose, for instance, that BA_1=0.45, BA_2=0.48 and BA_3=0.42, then BA_3 values will be corrected to 0.48. The upper bound of the 95th percentile (or higher if required) of the testing $BA_{1,2,...,h}$ values from the n samples are used to derive the 0.05 significance level (or lower if required). MDR is then run in the studied dataset and the top models/dimension are sorted out. The testing BA value for the 1st model is then compared to the significance threshold for the testing BA_1 values; if values are higher the model is declared significant, non-significant otherwise. In the latter case, the scaling procedure is applied and the testing BA values for the 2nd model is compared to the significance threshold for the testing BA_2 values. The procedure is repeated until a significant model is found up the hth model.

The power of the sMDR algorithm was tested via simulation in synthetic datasets. Since we are also aiming at determining the power of the sMDR algorithm in the systemic sclerosis (SSc) population, a few constrains were applied to simulation. Firstly, with a try-and-error approach we generated different penetrance functions for a triplet of biallelic genetic loci, with minor allele frequencies equal to 0.1, 0.2 or 0.3 (Table 2.1-10). The broad-sense heritability (H^2) of all the model was set to a relatively low value (=0.10); the environmental contribution to SSc pathogenesis is indeed thought to be low and we do not expect H^2 to be any lower. For each epistatic model, a population of 65000 individuals with the case-control status linked to the causative triplet was built. We then incorporated linkage disequilibrium (LD) in data simulation. The 2nd SNP of the triplet was linked to a single SNP with the same MAF (r^2=0.8) whilst the 3rd SNP of the triplet was linked to 2 SNPs with the same MAF (r^2=0.8). To these 6 SNPs (the triplet and the 3 linkage SNPs), 14 additional loci unlinked to the case-control-status with MAF ranging from 0.05 to 0.5 were added. Among these 14 SNPs, we considered 2 LD blocks of 2 SNPs and 3 LD blocks of 3 SNPs each (r^2=0.8). From this population we then draw 100 samples with 500 controls and 125 cases (case to control ratio = 1:4) or 500 controls and 250 cases (1:2) or 500 controls and 375 cases (1:1.25). These population would simulate the control vs diffuse SSc (dcSSc) or the control vs

limited SSc (lcSSc) or the control vs SSc comparison, where the dcSSc:lcSSc ratio is 1:2. Null distributions were then generated where instead of the triplet and of the triplet-linked loci, we considered an additional single locus and two additional LD blocks with 2 and 3 SNPs, respectively. One-hundred null samples were drawn from the null distribution with a case:control ratio as above and 0.05 or 0.025 thresholds for significance were derived. The scaling parameters (h) was set to 10 and 5 cross-validations were used. Specific power was estimated as the number of times sMDR identified the causative triplet of SNP out of 100 dataset/model/ case:control ratio at the desired significance threshold. A wider-sense power was estimated incorporating LD; in this case a positive hit was considered whenever the sMDR method sorted out the causative triplet or a triplet containing one of the loci in linkage with the triplet.

Simulation results depicted in table 1 clearly demonstrate that: a) sMDR retains a higher power than plain MDR to correctly identify the causative triplet of loci in the simulated datasets; b) due to a “ceiling” effect, the increase in power is more pronounced for models where MDR performs poorly (e.g. in the dcSSc vs controls comparison or when MAF are low); c) Incorporation of LD increases the power of both the MDR and the sMDR algorithm.

Overall these results indicate that sMDR is powerful enough to detect a causative triplet of loci or a set of loci in LD with the causative triplet in the SSc dataset we are currently analyzing.

Application of dMDR to the SSc dataset

To determine the thresholds for significance, the SSc/controls population as well as the lcSSc/controls and dcSSc/controls populations were was randomly permuted to generate 100 null datasets/population. Five cross-validations were used and the scaling parameter h was set to 10 and the top 95^{th} percentile of the testing BA_h values were derived from the null SSc/controls distribution, whilst for the null lcSSc/ controls and null dcSSc/controls distributions the top 97.5^{th} percentile were derived. Alpha level were thus set to 0.05 in the former case and, to account for multiple tastings, to 0.025 in the latter. The sMDR method was then applied to the populations under analysis considering the abovementioned significance levels.

Table 1 - Power for the Multifactor Dimensionality reduction (MDR) algorithm in the epistatic 3-factor models employed for simulation

		Power							
		Plain MDR		Scaling MDR		Plain MDR + LD		Scaling MDR + LD	
Model #	Cases:Controls	0.05	0.025	0.05	0.025	0.05	0.025	0.05	0.025
1	1:4	36	36	48	46	49	48	55	53
	1:2	76	76	93	93	93	92	98	98
	1:1.25	81	81	98	98	98	98	100	100
2	1:4	38	37	56	53	56	55	67	64
	1:2	95	94	97	95	95	94	97	95
	1:1.25	98	98	100	99	97	96	100	100
3	1:4	50	50	68	68	69	69	75	74
	1:2	83	83	96	96	98	98	100	100
	1:1.25	83	83	100	100	99	99	100	100
4	1:4	39	38	50	48	51	50	54	52
	1:2	76	73	86	83	89	85	87	84
	1:1.25	83	82	99	97	98	97	100	99
5	1:4	52	51	58	56	61	60	64	62
	1:2	91	91	99	98	96	96	100	100
	1:1.25	94	94	100	100	98	98	100	100
6	1:4	55	55	64	64	68	68	70	70
	1:2	91	91	98	98	97	97	100	100
	1:1.25	94	94	100	100	98	98	100	100
7	1:4	83	83	97	96	95	95	99	98
	1:2	97	95	99	99	98	98	99	99
	1:1.25	98	98	100	100	100	100	100	100
8	1:4	84	84	95	95	91	91	96	96
	1:2	98	98	100	100	99	99	100	100
	1:1.25	98	98	100	100	99	98	100	100
9	1:4	82	81	92	89	91	90	96	94
	1:2	98	98	100	100	99	99	100	100
	1:1.25	97	97	100	100	99	99	100	100
10	1:4	41	41	57	57	53	52	62	61
	1:2	67	66	91	90	88	87	95	94
	1:1.25	89	89	98	98	98	98	100	100

Table 2.1-10 – Penetrance functions for the different models employed for simulation

Model 1 – MAF (A) = 0.1; MAF (B) = 0.1; MAF (C) = 0.1

	AA			Aa			aa		
	BB	Bb	bb	BB	Bb	bb	BB	Bb	bb
CC	0,56775	0,82226	0,58733	0,78310	0,07831	0,70479	0,70479	0,46986	0,97888
Cc	0,74395	0,29366	0,58733	0,43071	0,72437	0,97888	0,46986	0,23493	0,46986
cc	0,78310	0,23493	0,03916	0,45028	0,19578	0,09789	0,70479	0,19578	0,93972

Model 2 – MAF (A) = 0.2; MAF (B) = 0.1; MAF (C) = 0.1

	AA			Aa			aa		
	BB	Bb	bb	BB	Bb	bb	BB	Bb	bb
CC	0,49332	0,75343	0,53816	0,69064	0,07176	0,53816	0,61889	0,43053	0,89694
Cc	0,60992	0,66374	0,48435	0,34084	0,89694	0,89694	0,43053	0,35878	0,43053
cc	0,69961	0,26908	0,03588	0,41259	0,71755	0,08969	0,17939	0,35878	0,86106

Model 3 – MAF (A) = 0.3; MAF (B) = 0.1; MAF (C) = 0.1

	AA			Aa			aa		
	BB	Bb	bb	BB	Bb	bb	BB	Bb	bb
CC	0,32223	0,60243	0,35025	0,50436	0,04203	0,35025	0,43431	0,33624	0,63045
Cc	0,47634	0,57441	0,44832	0,15411	0,91065	0,77055	0,33624	0,29421	0,40629
cc	0,54639	0,30822	0,09807	0,26619	0,70050	0,35025	0,02802	0,28020	0,63045

Model 4 – MAF (A) = 0.1; MAF (B) = 0.2; MAF (C) = 0.2

	AA			Aa			aa		
	BB	Bb	bb	BB	Bb	bb	BB	Bb	bb
CC	0,45657	0,66410	0,49807	0,83012	0,08301	0,68485	0,58108	0,45657	0,93389
Cc	0,58108	0,51883	0,49807	0,31130	0,62259	0,95464	0,64334	0,31130	0,47732
cc	0,76786	0,14527	0,08301	0,45657	0,64334	0,68485	0,24904	0,20753	0,97539

Model 5 – MAF (A) = 0.2; MAF (B) = 0.2; MAF (C) = 0.2

	AA	Aa	aa

	BB	Bb	bb	BB	Bb	bb	BB	Bb	bb
CC	0,28569	0,45451	0,23375	0,53243	0,01299	0,35062	0,44152	0,31166	0,59736
Cc	0,36361	0,37659	0,28569	0,14285	0,61034	0,92201	0,22076	0,31166	0,45451
cc	0,48048	0,11687	0,03896	0,25972	0,62333	0,38958	0,09090	0,12986	0,12986

Model 6 – MAF (A) = 0.3; MAF (B) = 0.2; MAF (C) = 0.2

	AA			Aa			aa		
	BB	**Bb**	**bb**	**BB**	**Bb**	**bb**	**BB**	**Bb**	**bb**
CC	0,2574	0,4095	0,2223	0,4563	0,0351	0,3159	0,3978	0,2925	0,5382
Cc	0,3276	0,3627	0,2574	0,1287	0,6435	0,5967	0,1989	0,2808	0,2574
cc	0,4329	0,1638	0,0351	0,234	0,5616	0,2808	0,0819	0,234	0,5265

Model 7 – MAF (A) = 0.1; MAF (B) = 0.3; MAF (C) = 0.3

	AA			Aa			aa		
	BB	**Bb**	**bb**	**BB**	**Bb**	**bb**	**BB**	**Bb**	**bb**
CC	0,12075	0,43168	0,36225	0,41113	0,04725	0,37001	0,37001	0,24668	0,51391
Cc	0,38850	0,15750	0,07350	0,22612	0,37275	0,50925	0,23625	0,12075	0,23625
cc	0,41113	0,13125	0,42000	0,23640	0,10278	0,05139	0,47250	0,10278	0,63000

Model 8 – MAF (A) = 0.2; MAF (B) = 0.3; MAF (C) = 0.3

	AA			Aa			aa		
	BB	**Bb**	**bb**	**BB**	**Bb**	**bb**	**BB**	**Bb**	**bb**
CC	0,11303	0,45214	0,33910	0,40076	0,08221	0,34938	0,38020	0,25690	0,51379
Cc	0,39048	0,16441	0,04110	0,20552	0,39048	0,46241	0,23634	0,10276	0,20552
cc	0,42247	0,16441	0,41103	0,24662	0,15414	0,04110	0,49324	0,15414	0,65765

Model 9 – MAF (A) = 0.3; MAF (B) = 0.3; MAF (C) = 0.3

	AA			Aa			aa		
	BB	**Bb**	**bb**	**BB**	**Bb**	**bb**	**BB**	**Bb**	**bb**
CC	0,12640	0,50560	0,36771	0,41368	0,12640	0,36771	0,43666	0,28728	0,57455
Cc	0,43666	0,18386	0,04596	0,22982	0,43666	0,49411	0,26429	0,11491	0,22982
cc	0,47113	0,24131	0,21833	0,27578	0,22982	0,03447	0,51710	0,17237	0,74692

Model 10 – MAF (A) = 0.3; MAF (B) = 0.2; MAF (C) = 0.1

	AA			Aa			aa		
	BB	**Bb**	**bb**	**BB**	**Bb**	**bb**	**BB**	**Bb**	**bb**
CC	0,23625	0,46125	0,25875	0,43875	0,06750	0,31500	0,35438	0,28125	0,51750
Cc	0,31500	0,38250	0,29250	0,09000	0,70875	0,58500	0,18000	0,27000	0,24750
cc	0,41625	0,22500	0,05625	0,21375	0,56250	0,28125	0,02250	0,22500	0,50625

References for the supplementary notes

Greene, C.S., Himmelstein, D.S., Nelson, H.H., Kelsey, K.T., Williams, S.M., Andrew, A.S., Karagas, M.R., Moorem J.H. (2010) Enabling personal genomics with an explicit test of epistasis. Pac Symp Biocomput. 327-36.

Moore, J.H., Gilbert, J.C., Tsai, C.T., Chiang, F.T., Holden, T., Barney, N., White B,C. (2006) A flexible computational framework for detecting, characterizing, and interpreting statistical patterns of epistasis in genetic studies of human disease susceptibility. J Theor Biol. 241:252-61.

Ritchie, M.D., Hahn, L.W., Roodi, N., Bailey, L.R., Dupont, W.D., Parl, F.F., Moore, J.H. (2001) Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. Am J Hum Genet. 69,138-47.

Part III

Epigenetics

Paul Klee, *Grenze*, 1938

Chapter 10

Skewed X-chromosomal inactivation impacts T regulatory cell function in systemic sclerosis

Broen JCA, Wolvers-Tettero ILM, Geurts-van Bon L, Vonk MC, Coenen MJH, Lafyatis R, Radstake TRDJ* and Langerak AW*

Annals of the Rheumatic Diseases 2011

Abstract

Objectives: To investigate the role of X-chromosomal inactivation (XCI) in systemic sclerosis and its effects on Foxp3 expression in T regulatory cells.

Methods: A total of 217 female SSc patients and 107 female healthy controls were included. From these subjects DNA was isolated from total PBMCs, plasmacytoid dendritic cells, T cells, B cells, myeloid dendritic cells and monocytes after magnetic bead separation. All samples were assessed for skewed XCI patterns with the HUMARA assay. Outcome was assessed with linear regression. Next, CD4+CD25+ cells were isolated and intracellular Foxp3 expression was assessed by flowcytometry.

Results: Skewing is not associated with increased age in SSc, in contrast to the control population (r=0.45, p<0.0001). Taking this into account, we found a significantly higher frequency of skewed XCI in SSc patients (P=0.001) compared to controls. No difference in skewing was observed between the immune cell subsets. In addition, we found a higher concentration of $Foxp3^+$ cells, exhibiting a lower Foxp3 MFI in the SSc patients with profound XCI skewing (both p<0.001), associated with less efficient suppressive activity (p=0.012).

Conclusions: Skewed XCI plays a role in susceptibility to SSc, is not restricted and influences Foxp3 expression and suppressive capacity of Tregs.

Introduction

Systemic sclerosis (SSc) is a severe autoimmune disease that is characterized by extensive fibrosis of the skin and internal organs culminating in high morbidity and premature death. Females develop SSc more often than males, reported ratios range from 5:1 up to 12:1 (1). Intriguingly, it was suggested that women with SSc contain a significantly higher frequency of peripheral blood cells with skewed X-chromosomal inactivation (XCI) patterns compared to their healthy counterparts (2,3). XCI is a dosage compensation mechanism evolved to level differences between males and females in X-chromosomal gene expression. As a consequence of this, females are functional mosaics for inactivation of the paternal or maternal X-chromosome (4). First, it was thought that X-chromosomal inactivation (XCI) encompasses the inactivation of all genes on one X-chromosome. However, a recent landmark study destabilized this paradigm by showing that at least 15% of the genes can escape XCI in physiologic conditions and are therefore expressed on two X chromosomes. In addition, heterogeneity was observed in the patterns of gene inactivation among females (5). Although inactivation of the X-chromosome is apparent to be permanent for all descendants of a cell, the XCI pattern alters with age. The frequency of skewed XCI in peripheral blood cells increases in elderly compared to younger healthy females (6). No studies investigated the potential functional relevance of XCI skewing in SSc, despite the fact that key immunomodulatory genes map to the X-chromosome. One of these concerns *forkhead box p3 (Foxp3)*, which is essential for regulatory T cells (Tregs) (7,8) Based on our recent study demonstrating markedly diminished suppressive capacity of Tregs in SSc (9), we here investigated whether XCI is present in a large cohort of well-documented SSc patients and to identify whether a skewed XCI could explain the aberrant Treg function.

Study design

Study population

A total of 217 female SSc patients and 107 female healthy controls were included, originating from the Boston University Medical Center in Boston, Massachussets, USA and the Radboud University Nijmegen Medical Centre in Nijmegen, The Netherlands. All patients met the American College of Rheumatology preliminary criteria for the classification of SSc (10). Both patients and controls were included in the study after written informed consent. The study was approved by the local ethical committees / institutional review boards. The mean age of the patients was 50 years (SD 14.5) and 46 years for healthy controls (SD 19.5).

Evaluation of XCI status utilizing the HUMARA assay

From each subject 200ng of DNA was used to evaluate XCI status according to methods described previously (11). Briefly, peak height of the shorter allele was divided by peak height of the longer allele and compensated for amplification bias using undigested controls. The skewing ratio was inverted if necessary to obtain a value of ≥1. A cell population with random XCI would be expected to show ratios equal or close to 1, which is an even distribution of 50% for each chromosome. For each person the ratio between the two chromosomes was calculated and the corresponding percentages are used throughout this article.

Isolation of cell subsets, assessment of FoxP3 expression and suppressive capacity

From 25 SSc patients and 9 controls we isolated PBMCs from heparinized venous blood by density-gradient centrifugation. Subsequently, BDCA4+, CD3+, CD19+, CD1C+ and CD14+ cells were isolated by magnetic cell separation techniques according to the manufacturer's protocol. DNA from all cell-types was extracted. After isolation the CD3+ fraction was stained with phycoerythrin, allophycocyanin and fluorescein isothiocynate conjugated mouse monoclonal antibodies against human CD4, CD25 and CD127. Intracellular staining of CD4+CD25+ cells for FoxP3 was performed according to the manufacturer (Miltenyi Biotec, Inc.) Assessment of expression was performed with a LSRII FACScan flow cytometer and emerging data were processed with FlowJo software. To assess Treg cell

suppressive capacity, $CD25^{high}CD127^{low}$ cells were sorted. $CD25^{high}CD127^{low}$ cells and unsorted CD3+ T cells were transferred into RPMI 1640 media supplemented with; 2nM L-glutamine, 100 U/ml penicillin-streptomycin, and 10% FCS. To assess the suppressive capacity of Tregs ($CD25^{high}CD127^{low}$), the unsorted T cells were brought to a concentration of $2x10^6$ cells/ml and stimulated with phytohaemaglutin. Tregs from healthy controls and SSc patients were added to autologous unsorted CD3+cells at fixed ratios 1:20 for 5 successive days. After 4 days of culture, [^{3}H]Tdr was added for the remaining 24 hrs of cultures. The cells were harvested onto glass fiber filters and [^{3}H]thymidine incorporation was assessed on a beta scintillation counter. The ability of Tregs to suppress proliferation was determined as the percentage of inhibition calculated as 1- (experimental counts per minute/ control counts per minute) x 100%.

Statistical analysis

Significance levels were calculated with linear regression taking into account the variables; age, the interaction between SSc-control status and age and affected status or the clinical phenotype of interest. Furthermore, Pearson correlations and Students t-test were calculated using SPSS 18.0 software. *P* values below 0.05 were considered as statistically significant. Values are displayed as mean plus/ minus SEM.

Results

Firstly, we investigated whether age has an effect on the degree of skewing observed in both the SSc patients and controls. We observed a clear association with increasing age in the control population (r=0.45, p<0.0001), whereas this correlation was not present in the SSc population, this resulted in different association results per age group (**Figure 1**).

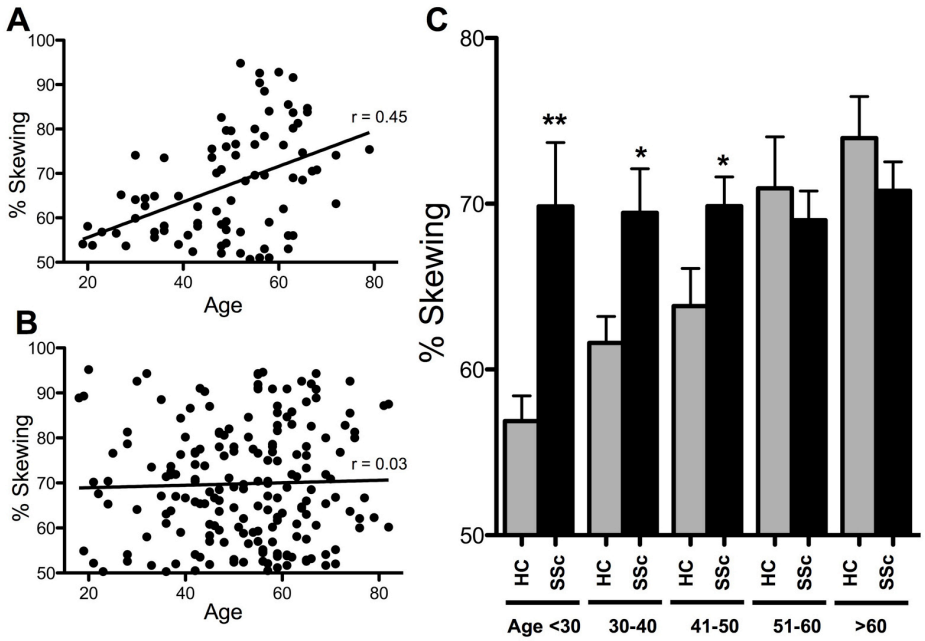


Figure 1: Correlation between age and skewing. A) There is a significant correlation between age and the degree of skewing present in the control population (r=0.45, P<0.0001) B) No correlation is present in the SSc population. C) Comparison of skewing between SSc and controls in different age groups, associations with skewing and SSc are present in subjects <30 years of age (P=0.006), 31-40 years (P=0.016) and 41-50 years (P=0.042) but are not significantly different between cases and controls after 50 years of age.

For this reason, we corrected for age effects with a linear regression model. In line with previous reports, we found a significantly higher frequency of skewed XCI in SSc patients (p=0.001) , lcSSc patients (p=0.008), dcSSc patients (p=0.005) and anti-topoisomerase autoantibody positive patients (p=0.008) compared to controls (**Table 1**). Since a skewed XCI pattern could reflect expansion of a subset of immune cells, we next evaluated the skewed XCI pattern in purified T cells (CD3+), B cells (CD19+), monocytes (CD14+), plasmacytoid dendritic cell (BDCA+), and

myeloid dendritic cells (CD1c+) from SSc patients and healthy controls (purity >98%, data not shown). All purified cell subsets showed a pattern of skewing similar to that seen in total PBMCs without any statistical differences.

Table 1. Clinical characteristics of patients included in this study and their association with skewing percentages by multiple linear regression, taking into account age effects.

Phenotype	n	Proportion of total	*P* *
SSc	192		**0.001**
lcSSc	94	0.49	**0.008**
dcSSc	98	0.51	**0.005**
Controls	84		
ACA+	43	0.22	0.260
Anti-Topo +	47	0.24	**0.008**
PAH+	39	0.20	0.295
Pulmonary Fibrosis +	69	0.36	0.266
Methotrexate	31	0.16	0.394
Azathioprine	31	0.16	0.783
Cyclophosphamide	23	0.12	0.179
Prednisolone	75	0.39	0.777

*All comparisons are performed with controls except PAH and Pulmonary Fibrosis, these were compared between patients with and without these clinical hallmarks.

XCI has functional consequences in that we demonstrated a clear association between skewed XCI (defined as a skewing ratio > 2 or 67%) and a lower mean fluorescence intensity of Foxp3 in $CD4^{+}CD25^{high}CD127^{-}$ compared to the patients with no skewing (P <0.001), in addition patients with no skewing have lower Foxp3 percentage positive cells compared to controls (P<0.001). Intriguingly, a higher percentage of Foxp3 (P<0.001) positive cells was observed in skewed SSc cases. To test the suppressive ability of Tregs in association with XCI, we studied the capacity of $CD25^{high}CD127^{-}$ cells from subjects with clear skewing (n=7) (skewing ratio > 2 or 67%), weak skewing (n=3) (defined as a skewing ratio between 55% and 67%) and no skewing (n=5) (defined as a skewing ratio <1.10 or 52%) to suppress the proliferation of CD4+ effector cells.(**Figure 2**).

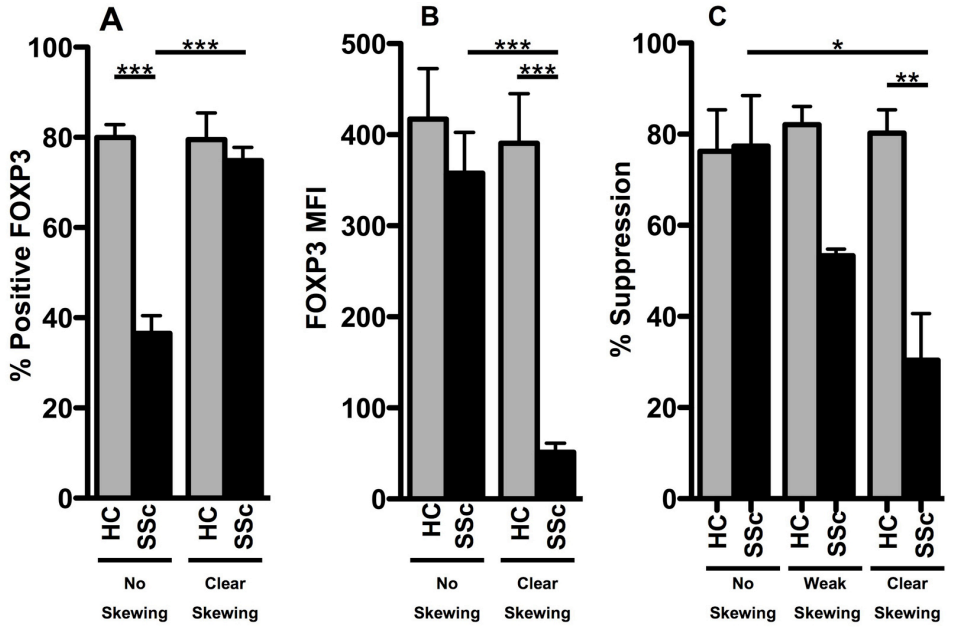


Figure 2: Effect of XCI skewing on $CD4^{+}CD25^{high}CD127^{-}$ cells. A) Percentage of FoxP3 positive $CD4^{+}CD25^{high}CD127^{-}$ cells in SSc patients with clear skewing (n=10) compared to patients (n=12) and controls (n=6) with no skewing (Both P<0.001). In addition, we observed a significant difference in the percentage positive in the controls with no skewing compared to the SSc patients with no skewing (n=4). B) Mean fluorescence intensity (MFI) of Foxp3 is different in patients with skewed XCI (n=10) compared to patients without skewing (n=12) and to healthy controls with skewing (n=4) (Both P<0.0001) no differences were observed between patients and controls (n=6) with no skewing. B) C). Suppressive capacity of $CD4^{+}CD25^{high}CD127^{-}$ cells from SSc patients and controls without XCI skewing (SSc n=5, HC n=4), clear XCI skewing (SSc n=7, HC n=4) and weak skewing (SSc n=3, HC n=2). Significant differences are present in the patients with skewing compared to the healthy controls and patients with no skewing (respectively P=0.007 and P=0.012).

The Treg populations in the no skewing, weak skewing and healthy control group had a similar suppressive capacity, whereas SSc patients displaying a clear skewing pattern exhibited a markedly diminished suppressive capacity of 30 % ± 10 compared to patients with no skewing (p=0.012) and healthy controls (p=0.007) **Figure 2C**). No significant differences in the number of surviving Tregs for each skewing and control group were observed after finishing the experiment, excluding biased results caused by early in vitro death of Tregs with skewed XCI.

Discussion

In this study we show that skewed XCI patterns are associated with SSc. In line with this, we demonstrate that XCI in SSc is not present in a specific cell subset, suggesting that this phenomenon is already present in precursor cells. Besides, our data firmly underscore the functional relevance of XCI in SSc by the strong association with a deregulated Treg capacity. There has been an ongoing discussion on the validity of the HUMARA assay in the determination of clonality in the past years (12). However, this discussion mainly focused on the ability of the HUMARA assay to detect clonality in elderly females. In the present study the vast majority of subjects is younger than 65 years and the results presented reflect the degree of skewing on a continuous scale without strictly assigning the subjects in groups with and without clonal expansion. Therefore this criticism does not apply to our findings (13). Important to note is that whenever one of the X-chromosomes is inactivated, both the *human androgen receptor* gene and *FOXP3* on this chromosome are silenced as well (5).

Human populations are outbred and differences between two X chromosomes are consequently present. It has been shown previously that these genetic differences influence hematopoietic stem cell kinetics, apparently leading to skewed XCI with aging (14). XCI skewing takes place in healthy females around the age of 55 years (15). Interestingly, the peak age at onset of SSc is close to this age, ranging from 49 to 64 years, depending on ethnic background (16). This is similar to that observed in other autoimmune diseases associated with skewed XCI patterns (17). We show that the increased XCI skewing in SSc takes place throughout all cells investigated, signifying that increased hematopoietic kinetics resulting from the autoimmune disease might accelerate the normal aging-associated process of XCI skewing.

However, this process appears to have a striking and unexpected effect on Treg function. Recently, it was shown that only Tregs and not other cell subsets selectively inactivate X-chromosomes carrying FOXP3 mutations (18). In contrast, we show that in SSc skewed XCI is associated with decreased FoxP3 expression. This suggests that altered X-linked genetic factors that influence hematopoietic

stem cell kinetics underlie the XCI skewing observed in SSc. These factors seem to prevail over the ability of Tregs to selectively inactivate X chromosomes that harbour regions less effective in cis and trans regulation of Foxp3. Recently, a study was published describing cis acting, non-coding, regulatory DNA elements in the Foxp3 gene and their impact on Treg fate. Intriguingly, mice deficient for one of these regulatory elements were able to induce expansion of $Foxp3^+$ cells as partial compensation for impaired Foxp3 induction (19). A similar mechanism could underlie our observations of decreased MFI and increased percentage of cells positive for Foxp3. Since peripheral blood cells have high regeneration rates, differences present between the two X chromosomes, might affect suppressive capacity of Tregs in the timeframe following skewing, making these females even more susceptible to autoimmune disease (20,21). Taken together, we present data that implicates XCI in the pathogenesis of SSc by deregulating FoxP3 expression, causing Treg dysfunction. It is tempting to speculate that XCI impacts on immune regulation in other autoimmune disorders as well.

References

1. Agarwal SK, Tan FK, Arnett FC. Genetics and genomic studies in scleroderma (systemic sclerosis). Rheum.Dis.Clin.North Am. 2008;34:17-40.

2. Ozbalkan Z, Bagislar S, Kiraz S et al. Skewed X chromosome inactivation in blood cells of women with scleroderma. Arthritis Rheum. 2005;52:1564-1570.

3. Uz E, Loubiere LS, Gadi VK et al. Skewed X-chromosome inactivation in scleroderma. Clin.Rev.Allergy Immunol. 2008;34:352-355.

4. Lyon, MF "Gene action in the X-chromosome of the mouse (Mus musculus L.)." Nature 190 (1961): 372-73.

5. Carrel L, Willard HF. X-inactivation profile reveals extensive variability in X-linked gene expression in females. Nature 2005;434:400-404.

6. Kristiansen, M., et al. "Twin study of genetic and aging effects on X chromosome inactivation." Eur.J.Hum.Genet. 13.5 (2005): 599-606.

7. Hori S, Nomura T, Sakaguchi S. Control of regulatory T cell development by the transcription factor Foxp3. Science 2003;299:1057-1061.

8. Lyon MF, Peters J, Glenister PH, Ball S, Wright E. The scurfy mouse mutant has previously unrecognized hematological abnormalities and resembles Wiskott-Aldrich syndrome. Proc.Natl.Acad.Sci.U.S.A 1990;87:2433-2437.

9. Radstake TR, van BL, Broen J et al. Increased frequency and compromised function of T regulatory cells in systemic sclerosis (SSc) is related to a diminished CD69 and TGFbeta expression. PLoS.One. 2009;4:e5981.

10. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee Arthritis Rheum. 1980;23:581-590.

11. Boudewijns M, van Dongen JJ, Langerak AW. The human androgen receptor X-chromosome inactivation assay for clonality diagnostics of natural killer cell proliferations. J.Mol.Diagn. 2007;9:337-344.

12. Swierczek SI, et al. “Hematopoiesis is not clonal in healthy elderly women.” Blood.;112(8) (2008):3186-93.

13. Busque L et al. “Skewing of X-inactivation ratios in blood cells of aging women is confirmed by independent methodologies.” Blood. 113(15) (2009): 3472-4.

14. Christensen K., Kristiansen M. et al. X linked genetic factors regulate hematopoietic stem cell kinetics in females. Blood 2000: 95(7): 2449-51

15. Di Nunzio S, Cecconi M et al. Wild-type FOXP3 is selectively active in CD4+CD25(hi) regulatory T cells of healthy female carriers of different FOXP3 mutations. Blood .2009: 114 (19):4138-41

16. Gale RE, Wheadon H, Boulos P, Linch DC. Tissue specificity of X-chromosome inactivation patterns. Blood 1994;83:2899-2905.

17. Wan YY, Flavell RA. Regulatory T-cell functions are subverted and converted owing to attenuated Foxp3 expression. Nature 2007;445:766-770.

18. Orstavik KH. Skewed X inactivation in healthy individuals and in different diseases. Acta Paediatr.Suppl 2006;95:24-29.

19. Zheng Y, Josefowicz S, Chaudhry A et al.Role of conserved non-coding DNA elements in the Foxp3 gene in regulatory T-cell fate.Nature. 2010 Feb 11;463(7282):808-12. Epub 2010 Jan 13.

19. Chifflot H, Fautrel B et al. Incidence and prevalence of systemic sclerosis: a systematic literature review. Semin Arthritis Rheum 2008;Feb;37(4): 223-35

20. Chabchoub G, Uz E, Maalej A et al. Analysis of skewed X-chromosome inactivation in females with rheumatoid arthritis and autoimmune thyroid diseases. Arthritis Res.Ther. 2009;11:R106.

Chapter 11

Telomere length in systemic sclerosis is clinical subtype and cell specific

Broen JCA, McGlynn L, Radstake TRDJ and Shiels PG

Submitted

Abstract

Aim: To assess telomere length in peripheral blood and immune cell subsets of healthy controls, limited and diffuse systemic sclerosis patients.

Methods: Relative telomere length of 182 SSc patients and 100 healthy controls was assessed using quantitative PCR. From 25 SSc patients and 9 healthy controls B cells, T cells, monocytes, plasmacytoid- and myeloid dendritic cells were isolated for immune cell specific telomere analyses.

Results: We observed a significant age related telomere attrition in healthy controls (n=100) and lcSSc patients (n=104) (Both $p< 0.001$), but not in dcSSc patients. In the immune cell subset specific analysis we observed significant shorter telomeres in B cells and myeloid dendritic cells of both lcSSc and dcSSc patients (B-Cells $p=0.014$, $p=0.002$ & myDCs $p=0.019$, $p=0.004$ respectively). PDCs and T cells were significantly shorter in dcSSc patients only (p=0.001 and p=0.003 respectively).

Conclusions: Aberrances in telomere shortening are a feature of SSc, reflecting a difference in clinical subsets at the cellular level.

Introduction

Systemic sclerosis (SSc) is an autoimmune disease characterized by vasculopathy, immune system activation and fibrosis of the skin and internal organs, which affects predominantly females (1). Over the last years, scientific research focusing on the factors underlying SSc pathogenesis and susceptibility intensified but, untill now, both a cure and an unequivocal disease model remain to be elucidated. However, during the passed years important steps forward have been made in the understanding of genetic susceptibility risks for SSc development. SSc does not inherit in a Mendelian fashion, but appears multifactorial, with an onset later in life. This implies that the combined effect of multiple small genetic variations in concert with environmental exposures determines the risk to develop SSc (2). Recently, this fact was further documented by a genome wide association study in SSc, showing that there was not one genetic factor posing enough risk to be fully accountable for SSc development (3). Despite extensive efforts, the identification of environmental factors ultimately determining the development of SSc in a genetically susceptible host remains to be established (4). A currently expanding field of research in SSc, providing a novel echelon for investigation situated between environment and genetics, is the realm of epigenetics.

Among other epigenetic processes, telomere length aberrations have been implicated in SSc. Telomeres are specialized nucleoprotein complexes at the end of eukaryotic chromosomes. They comprise tandem TTAGGG repeat arrays bound to a variety of proteins with roles in chromosomal protection, nuclear attachment and replication. Telomeres function to cap the chromosome, preventing chromosomal fusions and the recognition of the chromosome end as a DNA break. Telomeres facilitate chromosomal attachment within the correct sub-cellular compartment and have a critical role in DNA replication. The proteinaceous component of the telomere helps maintaining its structural integrity and functions in sensing, signalling and repair of DNA damage (5). The length of telomeric DNA repeats shortens during the ageing of cultured fibroblasts, peripheral blood lymphocytes and colon epithelia, but shortening is also under polygenic and environmental influence (6,7). As a consequence, telomere length reflects the “miles on the clock” of a certain individual or cell type. The pivotal role of this

complex is the protection of chromosomes from DNA damage by initiating cellular senescence, or apoptosis (8).

Increased chromosomal damage has been repeatedly reported in SSc lymphocytes as well as fibroblasts, (9-14). Interestingly, two previous studies, prompted by the observation of increased chromosomal damage in SSc, addressed telomere length in SSc patients and came to ambiguous conclusions (15,16). However, both studies used dissimilar methodology and clinical disease subsets in their investigations. To better gauge these previous findings, our study aims to evaluate telomere attrition in an independent, clinically well-defined SSc cohort. Moreover, we aim to investigate whether there are differences in telomere attrition between different immune cell subsets in SSc.

Study design

Patients and Controls

The study population was composed of 182 SSc patients and 100 healthy controls matched by geographical region and age. The population consisted of two case-control sets of European ancestry; a UK cohort: 90 SSc patients and 46 controls; a Dutch cohort: 92 SSc patients and 54 controls. All the patients fulfilled the 1980 American College of Rheumatology (ACR) classification criteria for SSc (17). The local ethical committee from each center approved the study. Both patients and controls were included in the study after written informed consent. All patients included in this study were classified as having limited cutaneous (lcSSc) or diffuse cutaneous SSc (dcSSc) using the criteria postulated by LeRoy (18). Patients with scleroderma changes limited to the skin distal to elbows and/or knees, regardless of facial involvement, fulfill the definition for lcSSc. Those SSc patients with more proximal scleroderma skin changes were classified as having dcSSc.

DNA/RNA isolation from peripheral blood leukocytes and cell subsets

Peripheral blood samples (10 ml) collected in EDTA tubes were obtained from each patient and stored at -80°C before DNA isolation. Genomic DNA was extracted from leucocytes in peripheral venous blood according to standard protocols. From 25 SSc (12 dcSSc and 13 lcSSc) patients and 9 controls, derived from the Boston University Medical Center (Boston, MA, USA) leucocytes were isolated from heparinized venous blood by density-gradient centrifugation. Subsequently, BDCA4+ (plasmacytoid dendritic cells), CD3+ (T cell), CD19+ (B cell), CD1C+ (myeloid dendritic cells) and CD14+ (monocyte) cells were isolated by magnetic cell separation techniques according to the manufacturer's protocol and as described previously (19). RNA was purified by AllPrep DNA/RNA columns (Qiagen, Valencia, CA), and cDNA was synthesized by I-script (Bio-Rad, Hercules, CA).

Telomere length analysis

DNA was extracted from PBLs following standard procedures and telomere lengths in the DNA samples were determined by Q-PCR, following the method of Cawthon as described previously (18). Telomere length determination was performed using

a Roche Light Cycler LC480. Telomere length analyses were performed in triplicate for each sample, using a single-copy gene amplicon primer set (acidic ribosomal phosphoprotein, 36B4) and a telomere-specific amplicon primer set. Quality control parameters employed for the amplifications comprised using a cut off of 0.15 for the standard deviation (SD) of the threshold cycle (Ct) for sample replicates. At an SD above 0.15 the sample was reanalyzed.

Relative telomere length was estimated from Ct scores using the comparative Ct method after confirming that the telomere and control gene assays yielded similar amplification efficiencies. This method determines the ratio of telomere repeat copy number to single copy gene number (T/S) ratio in experimental samples relative to a control sample DNA. This normalised T/S ratio was used as the estimate of relative telomere length (Relative T/S). The inter-assay variation was assessed by comparing the relative telomere estimates (T/S ratio) estimates across assays for the positive controls, which were assayed on every assay plate.

Statistical analyisis

Statistical analysis was performed using SPSS version 15 (SPSS Inc, Chicago, Illinois). Pearsons correlations were performed to establish any relationships between telomere length and the various parameters such as age and gene expression levels. Telomere length comparison between the diffent cohorts, disease groups and cell subsets was achieved using the Mann-Whitney test. Telomere length was corrected for age using ANCOVA analysis. Correction for multiple testing was performed using the Bonferroni correction.

Results

PBLs from diffuse SSc patients do not display age-related telomere attrition

Telomere lengths were determined for the PBLs from both the Scottish and Dutch cohort. However, to limit the confounding factors of the differences in the socioeconomic backgrounds of both cohorts, analysis of telomere attrition was performed on the two cohorts separately. For both the Scottish and Dutch cohort, we observed no significant differences in average age-corrected telomere length between controls, SSc, lcSSc and dcSSc. In the Scottish cohort, both the healthy controls as well as the lcSSc patients displayed evidence of age-related telomere attrition, with increasing age, telomeres were observed to significantly shorten (controls p=0.009, lcSSc p<0.001). Conversely, there was no noticeable change in the telomere lengths of Scottish dcSSc patients over an approximately 50 year period. Similar patterns were observed for the Dutch cohort, again the lcSSc patients displayed telomere attrition with increasing age (p=0.03), whereas the dcSSc patients maintained telomere length with increasing age. However, it has to be noted that the Dutch healthy controls did not display age-related telomere attrition, though this is perhaps due to the relatively young and narrow age range of this particular cohort (median Age=38yrs, interquartile range 30-49yrs). In the combined analysis both controls and lcSSc showed significant telomere attrition with age in contrast to the dcSSc patients (both p<0.001) (**Figure 1**).

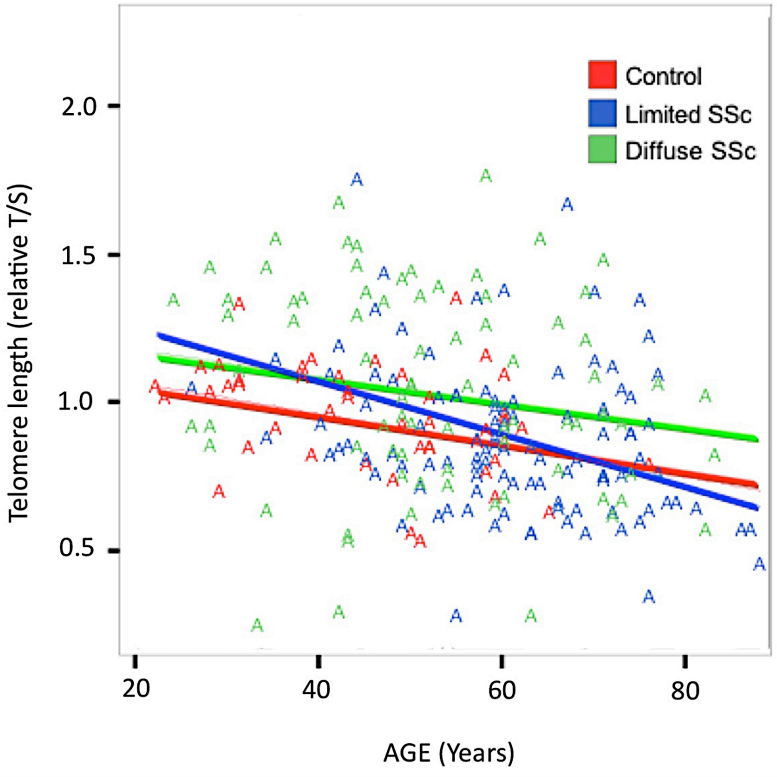


Figure 1: Significant age related telomere attrition in healthy controls (n=100) and lcSSc patients (n=104) (Both p< 0.001), but not in dcSSc patients.

Marked differences in telomere length of immune cell subsets from SSc patients

To address the question whether differences in telomere length are present between immune cell subsets in SSc, we isolated B cells, T cells, monocytes, myeloid dendritic cells and plasmacytoid dendritic cells. No significant differences in the age of the SSc, SSc subtypes and healthy controls were present.

In comparing healthy controls with SSc patients it was observed that for all the cell subsets, bar the monocoytes, telomere lengths were longer in the healthy controls. B-Cells (p=0.001, median relTS 1.15 vs 0.78), MDCs (p=0.008, median relTS 0.84 vs 0.62), PDCs (p=0.025, median relTS 0.84 vs 0.62) and T-Cells (p=0.004, median relTS 1.02 vs 0.78) from SSc patients had significantly shorter telomeres than the healthy controls. Interestingly for the B-Cell and myDCs this phenomenon seemed to be ubiquitously present in both the lcSSc and dcSSc subset (B-Cells p=0.014, p=0.002 and myDCs p=0.019, p=0.024 respectively). Conversely for both the PDCs and T-Cells we observed that these subsets presented with shorter telomeres in only the dcSSc population in comparison to the healthy controls (p=0.001 and p=0.003 respectively). The telomere lengths of PDCs and T-Cells from lcSSc patients were not significantly different from the controls. Evidence of dcSSc specific significant telomere shortening in PDCs (*P*=0.001) was strengthened by the determination that dcSSc PDCs had shorter telomeres than lcSSc PDCs (*P*=0.005). In contrast to the other cell subtypes, no significant differences were observed in the telomere length of monocytes between the different disease phenotypes (**Figure 2)**. However, after stringent correction for multiple testing using the Bonferroni corection; (5 cell subsets, three tests in each subset equals a corrected significance threshold of 0.05 divided by 15 tests is a treshold of p=0.0033), we have to conclude that only the observations of shorter telomeres in T cells, B cells and pDCs of dcSSc patients compared to healthy controls remain significant.

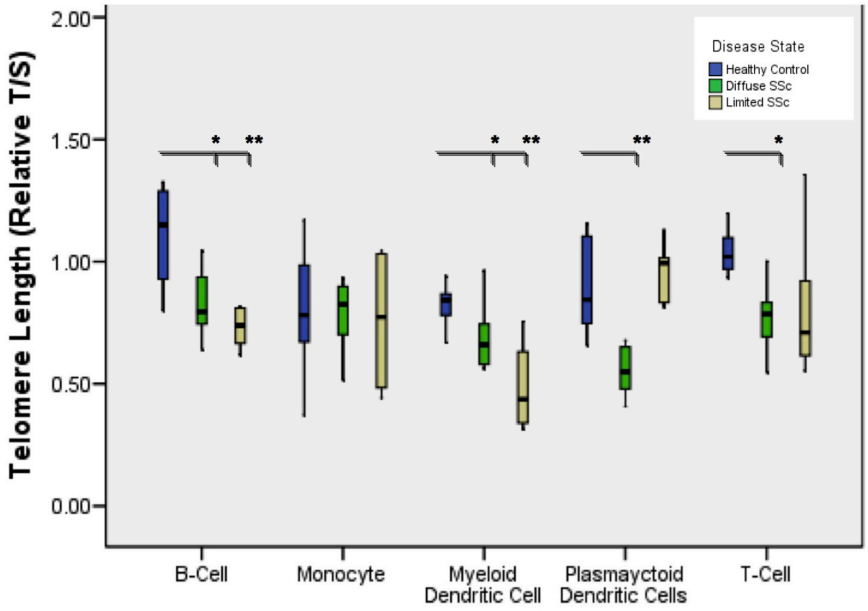


Figure 2. Significant shorter telomeres in B cells and myeloid dendritic cells of both lcSSc (n=13) and dcSSc (n=12) patients (B-Cells *p*=0.014, *p*=0.002 & myDCs *p*=0.019, *p*=0.004 respectively) compared to controls (n=9). PDCs and T cells were significantly shorter in dcSSc patients only (p=0.001 and p=0.003 respectively).

Discussion

In this study we describe an increased rate of age-related telomere shortening in lcSSc, in contrast to a decrease in dcSSc. Moreover, we show marked differences in telomere length in immune cell subsets of the clinical phenotypes of SSc. Two previous studies addressed telomere length in SSc. An initial study investigated telomere lengths of peripheral blood leukocytes (PBLs) and fibroblasts from 43 SSc patients, 182 SSc family members and 96 age-matched controls restriction fragment length polymorphism (RFLP) and chemiluminescent labelled probes. They observed an average loss of telomeric DNA in PBLs from SSc patients and their family members of 3 kb compared to the controls. This loss withstood correction for age and disease duration. Of interest, although telomeres in SSc fibroblasts were shorter overall compared to healthy control fibroblasts, this difference was not significant (15). This study used a different methodology than

our study and did not differentiate between lcSSc and dcSSc in their analysis, making it difficult to compare these results directly to our study. However, if they included particularly lcSSc patients at a higher age, these results would be well comparable with our results and those from a second study, showing increased telomere attrition in lcSSc patients, deviating from healthy controls at a higher age (15). The second study focused solely on females with the lcSSc phenotype. Forty-three lcSSc patients with an age ranging from 37 to 80 years were included. The difference between the lcSSc and control telomere lengths was only significant beyond the fifth decade. Below 50 years of age, no difference was observed between healthy females and females with lcSSc (16).

Considering our and previous observations it seems likely that telomeres shorten as a consequence of a higher turnover of PBLs and their precursor cells caused by the autoimmune disease SSc. However, there seems to be a marked difference between the immune cell subsets affected in dcSSc. In this clinical phenotype, PDCs and T cell seem to be more affected by telomere shortening compared to lcSSc. This might imply that these cells play a relatively larger role in dcSSc pathogenesis than in lcSSc. Although not significant after correction for multiple testing, B cells and myDC have shorter telomeres in both the limited and diffuse subform and the observed trend suggests involvement of these cells in both clinical phenotypes. Monocytes are not affected by telomere shortening in SSc, reflecting possibly no involvement or better counterbalancing mechanisms. Next to a higher demand on precursor cells, the decrease in telomere length of immune cell subsets might also be caused by a lack of telomere-repair gene or protein expression, such as telomerase, by the SSc immune cells. Another possibility is that mechanisms that trigger apoptosis upon reaching a telomere attrition treshold are not functioning adequately. This would keep immune cells with short telomeres circulating in the blood, decreasing average telomere length for these cells in our measurements. If this last possibility is true, it would also mean that there is a proportion of immune cells with a higher than normal content of DNA damage, more prone to derail into autoimmune reactivity (21).

In summary, for the first time in a large cohort, this study shows that there are marked differences between age related telomere attrition in clinical subforms of

SSc and that these differences are also reflected on the level of immune cell subsets. To better position these observations as a downstream or a causative effect, future investigations into the function of telomere maintaining genes and proteins are necessary, next to validation of these findings in a larger collection of immune cell subsets from SSc patients.

References

1. Varga J, Abraham D. Systemic sclerosis: a prototypic multisystem fibrotic disorder. J Clin Invest. 2007 Mar;117(3):557-67

2. Broen JCA, Coenen MJH, Radstake TRDJ. Deciphering the genetic background of systemic sclerosis. Expert Rev Clin Immunol. 2011

3. Radstake TR, Gorlova O, Rueda B, Martin JE, Alizadeh BZ, Palomino-Morales R, Coenen MJ, Vonk MC, Voskuyl AE, Schuerwegh AJ, Broen JC, van Riel PL, van 't Slot R, Italiaander A, Ophoff RA, Riemekasten G, Hunzelmann N, Simeon CP, Ortego-Centeno N, González-Gay MA, González-Escribano MF; Spanish Scleroderma Group, Airo P, van Laar J, Herrick A, Worthington J, Hesselstrand R, Smith V, de Keyser F, Houssiau F, Chee MM, Madhok R, Shiels P, Westhovens R, Kreuter A, Kiener H, de Baere E, Witte T, Padykov L, Klareskog L, Beretta L, Scorza R, Lie BA, Hoffmann-Vold AM, Carreira P, Varga J, Hinchcliff M, Gregersen PK, Lee AT, Ying J, Han Y, Weng SF, Amos CI, Wigley FM, Hummers L, Nelson JL, Agarwal SK, Assassi S, Gourh P, Tan FK, Koeleman BP, Arnett FC, Martin J, Mayes MD. Genome-wide association study of systemic sclerosis identifies CD247 as a new susceptibility locus. Nat Genet. 2010 May; 42(5):426-9. Epub 2010 Apr 11.

4. Mora GF. Systemic sclerosis: environmental factors. J Rheumatol. 2009 Nov;36(11): 2383-96. Epub 2009 Oct 1.

5. Lamb KJ, Shiels PG. Telomeres, ageing and oxidation. SEB Exp Biol Ser. 2009;62:117-37.

6. Carrero JJ, Stenvinkel P, Fellström B, Qureshi AR, Lamb K, Heimbürger O, Bárány P, Radhakrishnan K, Lindholm B, Soveri I, Nordfors L, Shiels PG. Telomere attrition is associated with inflammation, low fetuin-A levels and high mortality in prevalent haemodialysis patients. J Intern Med. 2008 Mar;263(3):302-12. Epub 2007 Dec 7.

7. Simpson RJ, Cosgrove C, Chee MM, McFarlin BK, Bartlett DB, Spielmann G, O'Connor DP, Pircher H, Shiels PG. Senescent phenotypes and telomere lengths of peripheral blood T-cells mobilized by acute exercise in humans. Exerc Immunol Rev. 2010;16:40-55.

8. O'Sullivan RJ, Karlseder J.Telomeres: protecting chromosomes against genome instability.Nat Rev Mol Cell Biol. 2010 Mar; 11(3):171-81. Epub 2010 Feb 3.

9. Migliore L, Bevilacqua C, Scarpato R. Cytogenetic study and FISH analysis in lymphocytes of systemic lupus erythematosus (SLE) and systemic sclerosis (SS) patients. Mutagenesis 1999;14:227–31.

10. Martins EP, Fuzzi HT, Kayser C, Alarcon RT, Rocha MG, Chauffaille ML, Andrade LE.Increased chromosome damage in systemic sclerosis skin fibroblasts.Scand J Rheumatol. 2010;39(5):398-401.

11. Housset E, Emerit I, Baulon A, de Grouchy YJ. Anomalies chromosomiques dans la sclérodermie: etude de 10 malades. Cr Acad Sci Paris 1969;296:413–16.

12. Wolff DJ, Needleman BW, Wasserman SS, Schwartz S. Spontaneous and clastogen induced chromosomal breakage in scleroderma. J Rheumatol 1991;18:837–40.

13. Pan SF, Rodnan GP, Deutsch M, Wald N. Chromosomal abnormalities in progressive systemic sclerosis (scleroderma) with consideration of radiation effects. J Lab Med 1975;86:300–8.

14. Porciello G, Scarpato R, Ferri C, Storino F, Cagetti F, Morozzi G, Spontaneous chromosome damage (micronuclei) in systemic sclerosis and Raynaud's phenomenon. J Rheumatol 2003;30:1244–7.

15. Artlett CM, Black CM, Briggs DC, Stevens CO, Welsh KI. Telomere reduction in scleroderma patients: a possible cause for chromosomal instability. Br J Rheumatol. 1996 Aug;35(8):732-7.

16. MacIntyre A, Brouilette SW, Lamb K, Radhakrishnan K, McGlynn L, Chee MM, Parkinson EK, Freeman D, Madhok R, Shiels PG. Association of increased telomere lengths in limited scleroderma, with a lack of age-related telomere erosion. Ann Rheum Dis. 2008 Dec;67(12):1780-2. Epub 2008 Jul 28.

17. Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. Preliminary criteria for the classification of systemic sclerosis

(scleroderma). Arthritis Rheum. 23(5), 581-590 (1980).

18. LeRoy EC, Black C, Fleischmajer R, Jablonska S, Krieg T, Medsger TA, Jr. et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. J Rheumatol 1988; 15(2):202-5.

19. Broen JC, Wolvers-Tettero IL, Geurts-van Bon L, Vonk MC, Coenen MJ, Lafyatis R, Radstake TR, Langerak AW. Skewed X chromosomal inactivation impacts T regulatory cell function in systemic sclerosis. Ann Rheum Dis. 2010 Dec; 69(12):2213-6. Epub 2010 Aug 10.

20. Cawthon RM. Telomere measurement by quantitative PCR. Nucleic Acids Res. 2002 May 15;30(10):e47.

21. Andrews NP, Fujii H, Goronzy JJ, Weyand CM. Telomeres and immunological diseases of aging. Gerontology. 2010 Jun;56(4): 390-403. Epub 2009 Dec 17.

Chapter 12

General discussion

General discussion

The aim of this thesis was to decipher part of the genetic background that underlies SSc susceptibility. Considering the manuscripts presented in this thesis, it can be concluded that this goal was reached. On the other hand it is still a small contribution in elucidating the role of genetics in SSc pathogenesis and susceptibility. Like in any other multifactorial disease, there is an extensive interplay between different genetic and environmental factors. Although this thesis scratches the surface of investigating interplay between genes in chapter 9, much work still needs to be done in this field. Prerequisites for such interaction studies are larger patient cohorts and dense genetic information. The studies described in this thesis, like all the other efforts in the area of genetic SSc research, evolved along with the revolution in genetic research possibilities. This project started with investigating single variants in candidate genes, (e.g. *STAT4, FAS, CD89*) moved to haplotype construction to perform whole candidate gene screening (*OX40L*) and subsequently to validating clinical relevance of SNPs and epigenetic influences.

At the point that genome wide screening methods came available, the use of simple candidate gene studies seemed to become redundant. Along with many others we observed that the main gaps in value of genetic research lie in the realm of functional validation and clinical translation, or freely put: “Functional validation is the bogeyman of genetic research” (according to Prof. J.A. Todd). Next to a genome wide association study (GWAS) being performed by our group in collaboration with many others, we started aiming at implementing these main omissions of genetic research in multifactorial disease within SSc genetic research. For this purpose a clinical follow-up database was constructed, to be able to predict the development of complications in time based on genetic markers. This is of particular use since the frequency of many complications in SSc increases based on disease duration. In addition, genotyping experiments were combined with multiple immune cell subset isolations derived from both healthy controls and SSc patients, encompassing B cells, T cells, plasmacytoid dendritic cells, monocytes and myeloid dendritic cells. Of these cell subsets both RNA and DNA were isolated, and they were used to investigate the functional consequences of the genetic observations in a cell specific matter. In addition ELISA was used to

better gauge the effect of variants on protein levels in the blood of patients. These efforts resulted in the findings described in part II of this thesis. For some variants we were able to indeed show an effect of genetic variants on susceptibility, complication development and functional impact (e.g. *TLR2, PBEF*), for others we did not observe a relevant effect (*CD89*, *IL4, IL13* and corresponding receptors). The approach of combining clinical follow-up data and functional implication in multiple immune cells allowed us to draw some more sustainable conclusions than the ones based on candidate gene studies alone. In addition, the use of follow-up data provided information about time related complication development, an area that was scarcely explored by SSc genetic research in a proper manner before. It is to be hoped that these studies will be able to form a foundation for a bridge across the gap of multifactorial genetics, functional relevance and clinical reality in SSc.

Another field that has rapidly increased during the last years is epigenetics. Methylation, microRNA regulation and age related DNA changes have become of high interest. These processes have been found implicated in SSc previously, as described in the introduction. In part III of this thesis previous findings of skewed X chromosome inactivation were validated in our cohort. In addition, it was investigated whether there were any differences between cell subsets, as well as if skewing affected regulation of genes located on the X chromosome. Although these findings are still quite preliminary in the sense that they provide little mechanistic data, they might prove to be important in the misbalance between males and females that are affected by SSc. To really pinpoint the changes that occur during this process, a databank with DNA, RNA and protein information on multiple time points, in multiple immunecells needs to be established. In this way changes in the amount of skewing can be pursued, as well as its impact on gene and protein expression in time. In addition, telomeres shortening in B cells and plasmacytoid dendritic cells from SSc patients, not corresponding to the declines associated with age is described in this thesis. This can be partially explained by increased proliferation of these immune cell subsets, but also indicates that these cells are not able to cope with the shortening correctly, either by going into apoptosis or by appropriately increasing telomerase expression to restore telomere

function. This epigenetic malfunctioning might contribute to the increased numbers of these cells observed in SSc.

Taken all these investigations together with other genetic research in SSc, there is a vast body of evidence implicating genetics in SSc pathogenesis. This data is still partially contradictory, which is probably attributable to the relatively small sample sizes used. This is a common problem in genetic research. Luckily, and despite SSc being a rare disease, multi-centre initiatives to investigate the role of genetics in SSc are increasing. This will ultimately lead to more homogenous results, although some rare variants might need very large populations (e.g. 10,000 SSc samples) before being uncovered as risk factor. One of the most interesting findings is that the genetic background of SSc seems to be very similar to that of other AID. This observation is partially biased because a large number of candidate genes are selected based on the principle of involvement in other AID. Interestingly, genetic variants associated with SLE seem to be directive for the candidate genes investigated in SSc, whereas RA seems to be the most prominent other concomitant autoimmune disease present in families of patients with SSc. Although RA has a higher prevalence than SLE, this might signify that genes involved in RA might also be good indicators for candidate studies in SSc. This is underlined by non-hypothesis driven approaches, in this case GWAS, which have yielded susceptibility genes found associated in SLE, RA and Sjögrens syndrome as well. An important gap in our understanding of the genetics involved in SSc is the lack of functional validation of genetic findings in immune cells or fibroblasts. This is becoming more apparent to researchers as there is a marked increase in studies that, next to describing genetic aberrances, aim at validating these results functionally. This will ultimately lead to a better understanding of the genetic component, but also identify functional pathways present in SSc pathogenesis. To date, the single most important step forward in the field of SSc genetics are the large GWAS recently published. These studies indicate that common disease variants are involved in SSc and strongly implicates the HLA region. Genome wide association studies are a powerful tool, but have shortcomings. The analysis of data is still very rudimentary because it assesses significance for each SNP separately. In a multifactorial disease, one would expect that clusters of genetic variants contribute together to the total genetic disease susceptibility. The logical

step would be to investigate these clusters and thus to identify those combinations of variants that cumulate to the total genetic susceptibility risk.This field, which is currently expanding, is the gene-gene interaction analysis, which is able to discover additive effects and might be able to explain a larger part of SSc disease risk than the current genetic markers alone. It is to be hoped that this will further evolve along with increasing computer power, required to deal with the high computational burden posed by these analyses.

Future directions

The cohorts of patients that are investigated for genetic risk factors in SSc have steadily increased over the last years and will keep on expanding. However, larger sample sizes are still necessary to identify common variants with small effects. For this purpose it is likely that even more global collaboration between SSc DNA databanks and rheumatologists will emerge in the upcoming years. The observation that common genetic variants are not the only factors contributing to SSc susceptibility, together with the similar genetic background present in other AID, indicates that rare variants, epigenetic changes or environmental triggers are able to determine which autoimmune disease will develop in a susceptible person. This opens the door in the next years for new technologies such as next generation full exome/genome sequencing, which is able to identify rare genetic variants, and DNA methylation profiling. Rare variants are likely to affect SSc susceptibility, but may differ between patients and are possibly de novo or evolutionary “young” mutations. Therefore, this technique will be especially powerful in families with SSc. Studies addressing methylation patterns are of high interest in SSc, since these are able to identify those regions of the genome that are part of the epigenetic inheritance involved in SSc pathogenesis. This would be extremely informative in a MZ twin study, since it might explain a proportion of the discordance observed in twins from SSc patients. Furthermore, studies that would shed light on the development of chromosomal aberrances in SSc might point to processes involved early in the disease onset. In addition a lot is expected from screening micro-RNA’s for post transcriptional regulation of gene expression. This novel level of gene expression regulation might be an important step in understanding the development of AID on a certain genetic background. Another

leap forward will be the transition to personal genomics. Next to traditional risk factors used in clinics for diagnostic and treatment processes, common genetic risk factors, rare variants and epigenetics will be used to determine risks for developing complications or response to treatment. In SSc it would be particularly interesting to identify genetic risk factors that play a role in development of scleroderma renal crisis, pulmonary hypertension and pulmonary fibrosis. When these variants are known, physicians are able to identify patients at risk early and start preventive treatment of these incurable and often fatal complications. Genetic research in the field of SSc will probably sail the winds of changes in genetic research methodology in the upcoming years, which hopefully will bring us to the brink of understanding this complex disease.

Summary

Systemic sclerosis (SSc) is a systemic autoimmune disease which particularly affects the connective tissue. The disease is rare with currently about 1600 patients in the Netherlands and an estimated incidence of 3 to 22.8 per million per year. Women are more often affected than men; the literature reports a range of 3-14 women against one man. Although more risk factors are identified, the pathogenesis of SSc remains largely unresolved.

The disease appears to be based on a triad of processes, vasculopathy, immune activation and fibrosis. The vasculopathy first manifests itself as Raynaud's phenomenon. This cold and stress triggered discoloration of the fingers and feet from white to blue to red is always present in SSc. On the other hand, only very few people with Raynaud's develop full-blown SSc. Exacerbation of the vasculopathy may result in ischemia; causing digital ulcers, necrosis and loss of the fingerpads. Manifestations of vasculopathy that occur later in the disease are renal crisis and pulmonary arterial hypertension. The second pillar of the disease process, immune activation, is characterized by the production of auto-antibodies, B-and T-cell activation and perivascular infiltrates. The current hypothesis is that these immune cells produce cytokines that incite fibroblasts to produce excessive collagen matrix, leading to fibrosis or sclerosis of the skin and internal organs. The diagnosis of SSc can be made by using the criteria of the American College of Rheumatology. If there is scleroderma (sclerosis of the skin) or if two of the following symptoms are present: 1) sclerodactyly, 2) digital ulcers or loss of fingertip, or, 3) bibasilar pulmonary fibrosis. Based on the pattern of skin sclerosis, a classification can be made in limited SSc (lcSSc) and diffuse SSc (dcSSc), both subforms have characteristic autoantibodies and complications. In more detail, lcSSc is characterized by a long-standing history of Raynaud's phenomenon. Fibrosis of the skin is present on the hands, face, feet and forearms. In addition, the presence of anti-centromere antibodies is very common (70-80%) and patients with lcSSc more often develop pulmonary hypertension. Diffuse SSc patients, usually suffered less than one year from Raynaud's before full-blown SSc unveils itself. In addition, interstitial pulmonary fibrosis, gastrointestinal fibrosis, renal failure and myocardial problems occur early in the disease process. The skin on

the acra and trunk is fibrotic. Thirty percent of dcSSc have anti-topoisomerase antibodies. Although the prognosis and treatment of SSc has improved in recent years, the disease remains incurable. Since complications for each clinical subform are different, the prognosis differs for lcSSc and dcSSc. Limited SSc has a 10-year survival of 75% -79%, diffuse SSc 53% -62%. However, when pulmonary hypertension develops in the lcSSc patients the survival drops to about 10% over 5 years.

This thesis is divided into three parts. The first part is based on four studies into genetic variants, each covers one chapter. In chapter 2, a variant of the *STAT4* gene is discussed, which is involved in the functioning of T cells. The variant is more common in patients with SSc and therefore plays a role in the susceptibility to developing the disease. In chapter 3, a genetic variant of *FAS* is scrutinized; this gene is involved in initiating regulated cell death. The variant is linked to the expression of FAS in cells. It appears that patients with positive anti-centromere antibody titers and limited disease carry this variant more often than healthy people. In the fourth section a variant in *CD89* is described; this variant plays no role in SSc and RA. The fifth chapter confirms a previously reported association of a polymorphism in the *TNFSF4* gene using a large European population of SSc patients.

The second part dives deeper into the role that genetic variants play in the functioning of immune cells and the development of clinical complications. In chapter 6 it is shown that various polymorphisms in the genes for interleukins 4 and 13, and their receptors, play no significant role in the development of SSc, associated complications and production of these cytokines. This is in contrast to the findings in chapter 7, where after an evaluation of various genetic variants in Toll like receptors, we show that a variant of the Toll like receptor 2 gene predisposes for developing diffuse SSc and pulmonary arterial hypertension (PAH). Furthermore, it is demonstrated that this variant, if present on dendritic cells, increases the release of inflammatory proteins. Chapter 8 starts with the observation that expression of the PBEF protein is higher in SSc patients with PAH. Next it is shown that this expression is influenced by two previously described variants in the promoter region of the PBEF gene. These variants also seem to

play a role in the development of PAH during the period of illness. The variant also seems to be involved in the susceptibility to develop SSc. In the final chapter of this section, we analyze a modified analysis method for interactions between different genes and variants (called epistasis) This pointed to a genetic interaction between a variant in the *PBEF* gene, *TLR9* and *Interleukin 1-alpha*, which could be validated at a functional level using cell systems.

The third part investigates processes that take place at a higher level of regulation than the DNA-sequence (epigenetics). Chapter 10 investigates the phenomenon of X chromosomal inactivation in SSc. Normally, women have two X chromosomes of which one is randomly inactivated. As women age, this random inactivation increasingly shifts to a skewed distribution. Chapter 10 shows that women with SSc display X skewing already early in life. Furthermore, the expression of a gene on the X chromosome, *Foxp3*, is affected in regulatory T cells. These cells play an important role in suppressing inflammatory responses, and seem to be rendered less capable doing so by this X chromosomal process. In the last chapter, the length of telomeres in immune cells from healthy controls and SSc patients was examined. Telomeres are the ends of chromosomes that shorten with each cell division and thus shorten in aging cells. In patients with SSc, telomeres shorten in a different fashion as compared to healthy individuals. Next to this, there are clear differences in telomere length between different immune cell types that are specifically for limited or diffuse SSc. This may reflect a difference in the involvement of these cells in different subtypes of SSc.

This thesis makes a modest contribution to the understanding of genetic processes that play a role in SSc. The challenge for the coming years lies in translating genetic findings into the clinic, as therapeutic targets and complication predictors.

Samenvatting

Systemische sclerose (SSc) is een systemische auto-immuunziekte die met name het bindweefsel aantast. De ziekte is erg zeldzaam met momenteel ongeveer 1600 patiënten in Nederland en een geschatte incidentie van 3-22,8 per miljoen per jaar. Vrouwen zijn vaker aangedaan dan mannen, de literatuur rapporteert een spreiding van 3-14 vrouwen tegenover 1 man. Hoewel er steeds meer, met name genetische, risicofactoren bekend worden blijft de pathogenese van SSc grotendeels onopgehelderd.

De ziekte lijkt te zijn gefundeerd op een trias van processen; vasculopathie (vaatafwijkingen), immuunactivatie en fibrose (verlittekening). De vasculopathie uit zich over het algemeen als eerste in de hoedanigheid van het fenomeen van Raynaud. Deze koude en stress gevoelige verkleuring van de vingers en tenen van wit naar blauw naar rood komt altijd voor bij SSc, maar andersom krijgen maar zeer weinig mensen met het fenomeen van Raynaud SSc. Bij verergering van de vasculopathie kunnen digitale ulcera (zweren aan de vingers), necrose (versterving) en verlies van de vingertoppen ontstaan door een tekort aan zuurstoftoevoer. Uitingen van vasculopathie die later in het ziekteproces optreden zijn pulmonale arteriële hypertensie en renale crise (ernstige long en nierafwijkingen). De tweede pijler van het ziekteproces is immuunsysteem activatie, dat zich kenmerkt door de productie van auto-anti-lichamen, B- en T-cel activatie en perivasculaire infiltraten. De huidige hypothese is dat deze immuuncellen cytokines produceren die fibroblasten aanzetten tot productie van overmatige collagene matrix, leidend tot fibrose oftewel sclerose van de huid en interne organen. De diagnose SSc is te stellen volgens de American College of Rheumatology wanneer er sprake is van sclerodermie (sclerosering van de huid) of wanneer twee van de volgende symptomen aanwezig zijn: 1) sclerodactylie, 2) digitale ulcera of verlies vingertop 3) bibasilaire pulmonale fibrose. Intrigerend is dat op basis van het patroon van huidsclerose een indeling gemaakt kan worden in gelimiteerde SSc (lcSSc) en diffuse SSc (dcSSc), met beide karakteristieke auto-antilichamen en complicaties. In meer detail; lcSSc kenmerkt zich door een anamnese van een al jaren bestaand Raynaud's fenomeen. Bij onderzoek is er sclerose van de huid aanwezig op handen, gezicht, voeten en onderarmen,

daarnaast zijn er vaak teleangectasien en huid calcificaties waar te nemen (calcinosis cutis). De aanwezigheid van anti-centromeer antilichamen bij laboratorium onderzoek komt zeer vaak voor (70-80%). Patiënten met lcSSc hebben vaker pulmonale hypertensie dan patiënten met diffuse SSc. Diffuse SSc patiënten hebben daarentegen korter dan 1 jaar last van het fenomeen van Raynaud. Daarnaast treedt vroeg in het ziekteproces interstitiële longfibrose, gastrointestinale fibrose, nierfalen en myocardiale problematiek op. De patiënt heeft met name last van kortademigheid, zuurbranden en verminderde urineproductie. Bij onderzoek van de huid bevindt zich de sclerose proximaal op het lichaam. Dertig procent van dcSSc hebben anti-topoisomerase antilichamen. Ongeveer 1/3 van alle SSc patienten heeft dcSSc en 2/3 lcSSc. Alhoewel de prognose en behandeling van SSc de laatste jaren verbeterd is, blijft de ziekte ongeneeslijk. Voortvloeiend uit de verschillende orgaancomplicaties passende bij de gelimiteerde en diffuse vorm van SSc is ook de prognose verschillend. Gelimiteerde SSc heeft een 10 jaars survival van 75%-79%, diffuse SSc echter 53%-62%. Ontstaat er echter pulmonale hypertensie gedurende het ziektebeloop van de gelimiteerde vorm, dan zakt de survival alsnog naar ongeveer 10% over 5 jaar.

Dit proefschrift is ingedeeld in drie onderdelen. Het eerste onderdeel beschrijft een viertal genetische varianten, behandeld per hoofdstuk. In hoofdstuk 2 wordt een variant van het *STAT4* gen besproken, een gen betrokken bij onder andere het functioneren van T cellen. Deze variant komt vaker voor bij patiënten met SSc en speelt daarom mogelijk een rol bij de gevoeligheid voor het ontwikkelen van de ziekte. In hoofdstuk 3 komt een genetische variant van *FAS* aan de orde, dit gen is betrokken bij gereguleerde celdood. Deze variant is in staat om de expressie van FAS in cellen te veranderen. Het blijkt dat patiënten met antilichamen tegen centromeren en gelimiteerde ziekte, vaker deze variant dragen dan gezonde mensen en andere SSc patiënten. In het vierde hoofdstuk wordt beschreven dat een variant in CD89 geen rol speelt in SSc. In het vijfde hoofdstuk wordt een eerdere associatie van een polymorfisme in het *TNFSF4* gen met SSc bevestigd in een grote populatie Europese SSc patienten.

Het tweede onderdeel van het proefschrift duikt dieper in de rol die genetische varianten spelen in het functioneren van immuuncellen en de ontwikkeling van klinische complicaties. In hoofdstuk 6 is beschreven dat een aantal polymorfismen in de genen voor *interleukines 4* en *13*, en hun receptoren, geen rol van betekenis spelen in het ontwikkelen van SSc en de bijpassende complicaties en productie van deze cytokines. Dit in tegenstelling tot de bevindingen in Hoofdstuk 7, alwaar na een evaluatie van genetische varianten in toll like receptors aangetoond wordt dat er een variant in het *TLR2* gen bestaat die een grotere kans geeft op het ontwikkelen van diffuse SSc en PAH. Daarnaast wordt aangetoond dat deze variant in dendritische cellen een toename geeft van ontsteking stimulerende eiwitten. Hoofdstuk 8 start met de observatie dat de expressie van het eiwit PBEF hoger is in SSc patienten met PAH, vervolgens wordt aangetoond dat deze expressie onder invloed staat van twee eerder beschreven varianten in de promoter regio van het *PBEF* gen. Deze varianten lijken ook een rol te spelen in de ontwikkeling van PAH gedurende de ziekteperiode. Daarnaast lijkt de variant een rol te spelen in de gevoeligheid voor ontwikkeling van SSc. In het laatste hoofdstuk van dit onderdeel wordt een nieuwe analysemethode toegepast om na te gaan of er interacties tussen verschillende genen en varianten bestaan (epistase genoemd). De analyse geeft aan dat er een genetische interactie is tussen een variant in het *PBEF* gen, *TLR9* en *Interleukine 1-alpha,* die ook *in vitro* lijkt te bestaan.

Het derde onderdeel ontstijgt het niveau van de DNA code en kijkt naar processen die daar net boven plaatsvinden. Normaal gesproken hebben vrouwen een tweetal X chromosomen waarvan er één “random” oftewel willekeurig wordt uitgeschakeld. Naarmate vrouwen ouder worden verschuift deze willekeurige uitschakeling naar een “skewed” oftewel een voorkeursverdeling voor één van de 2 X chromosomen. Hoofdstuk 10 laat zien dat vrouwen met SSc al vroeg in hun leven een voorkeursverdeling hebben. Daarnaast is de expressie van een gen op het X chromosoom, *FOXP3,* verlaagd in T regulatoire cellen. Deze cellen spelen een belangrijke rol in het onderdrukken van ontstekingsreacties, en lijken daar, door dit X chromosomale proces, minder goed toe in staat. In het laatste hoofdstuk wordt de lengte van telomeren in immuuncellen van gezonde controles en SSc patienten onderzocht. Telomeren zijn de uiteinden van chromosomen die bij iedere celdeling

verkorten en dus korter worden bij veroudering van cellen. In patienten met SSc verkorten deze telomeren niet zoals in gezonde personen, daarnaast zijn duidelijke verschillen aan te wijzen tussen verschillende soorten immuuncellen die specifiek in gelimiteerde of diffuse SSc verkorte telomeren hebben. Mogelijk weerspiegelt dit een verschil in betrokkenheid van deze cellen in de verschillende subtypes van SSc.

Dit proefschrift levert een bescheiden bijdrage aan het inzicht in genetische processen die een rol spelen bij SSc. De uitdaging voor de komende jaren ligt echter in de vertaling van genetische bevindingen naar de kliniek in de vorm van therapeutische aangrijpingspunten en complicatie beperkingen.

Dankwoord

Op de eerste plaats wil ik de patiënten en vrijwilligers bedanken die belangeloos hun DNA hebben afgestaan en daarmee de onmisbare basis voor het hier gepresenteerde onderzoek vormen.

Geachte professor van Riel, beste Piet. Hartelijk bedankt voor je ondersteuning en het in mij gestelde vertrouwen gedurende het doorlopen van dit promotietraject en het altijd klaar staan voor het wegnemen van obstakels en beantwoorden van vragen.

Geachte professor Radstake, beste Tim, inmiddels vier jaar geleden heb ik een eerste gesprek met je gehad toen ik solliciteerde voor de functie van student-assistent. Je gaf toen aan dat ik daarna een wetenschappelijke stage in de VS kon doen, gevolgd door promotieonderzoek en later misschien wel in opleiding voor de reumatologie. Klinkt fantastisch natuurlijk in de oren van een derdejaars geneeskundestudent, maar het leek mij toch wel een wat utopische voorstelling. Maar, eerlijk is eerlijk, je hebt gelijk gekregen. Dat was nooit gelukt zonder jouw nooit aflatende enthousiasme, geduld en begeleiding. Bedankt dat je me de kans hebt gegeven om me naast mijn medische opleiding wetenschappelijk te ontplooien en dat je me de kans geeft om me nu in Utrecht verder te ontwikkelen.

Geachte Dr. Coenen, beste Marieke. Mijn eerste wetenschappelijke stapjes heb ik gezet in het lab multifactoriële ziekten. Van jou heb ik geleerd hoe ik een pipet moest vasthouden en dat het genoom van *mus musculus* niet van de mus, maar van de muis afkomstig is. Belangrijker is echter dat je me geleerd hebt kritisch te zijn op methoden en resultaten en je de basis vormt voor de genetische interesse en kennis die ik nu heb. Bedankt voor je begeleiding en gastvrijheid de afgelopen jaren en ik hoop dat we ook in de toekomst kunnen blijven samenwerken.

Beste collega's van Groep Radstake, ik wil jullie hier persoonlijk bedanken:
Lenny, samen zijn we begonnen aan het promotie-avontuur in Boston, het was wetenschappelijk en sociaal een geweldige tijd die we daarna hebben kunnen voortzetten in Nijmegen en nu in Utrecht. Ik ben blij om je tot mijn collega's en

vrienden te mogen rekenen.
Mark en Kim, beiden collega's van het eerste uur, fijn dat ik bij jullie altijd terecht kan met vragen over immuuncellen, FACS en MACS. Verder hoop ik dat er nog veel gezellige bierproeverijen en spelavonden samen mogen volgen.
Marta, je bent een geweldige collega en ik beleef inmiddels dagelijks veel plezier aan je humor en frisse (Italiaanse) blik op zaken.
Richard en Tamara, bedankt voor de prettige samenwerking en de vele prikkelende wetenschappelijke discussies.

Beste collega's van lab multifactoriële ziekten, zonder jullie was dit proefschrift waarschijnlijk nooit verschenen. *Mascha*, bedankt voor je begeleiding en uitleg van technieken, het altijd klaar staan als er iets gedaan moest worden als ik met mijn coschappen bezig was, maar het belangrijkste was wel de gezellige, humorvolle dagelijkse samenwerking. *Remco,* bedankt voor de prettige samenwerking en al je hulp met automatische analyses. *Marlies* en *Angelien,* jullie positieve instelling en energie zijn belangrijke ingrediënten voor de goede sfeer in het lab, ik vond het fijn om met jullie samen te werken. *Johanne,* bedankt voor al je hulp met het SSc DNA. *Barbara*, bedankt voor de gastvrijheid in het lab en de waardevolle wetenschappelijke discussies tijdens werkbesprekingen.

Javier, Lara, Blanca, Ezequiel and the many other people involved in SSc genetics from Granada; the importance of your involvement in creating this thesis is beyond doubt. Many thanks for all your help and contributions to this thesis and the nice times during visits and congresses. I sincerely hope we will keep on collaborating in the future and that many social and scientific events may follow from it.

Dear Paul and colleagues from Glasgow, thanks for the opportunity to stay and learn in your lab and the possibility to enjoy the many intriguing faces of Glasgow together.

Dear Professor Lafyatis, dear Bob, thanks for providing the opportunity to perform a traineeship in your lab, my time in Boston has definetely ignited the will to obtain a PhD in scleroderma research.

Madelon en Hanneke, hoewel we elkaar weinig gesproken hebben, was een groot deel van dit proefschrift er niet gekomen zonder jullie bewonderenswaardige begaandheid met de systemische sclerose patiënten en de daaruit voortkomende wil bij te dragen aan wetenschappelijk onderzoek naar deze ziekte. Mijn hartelijke dank daarvoor.

Geachte leden van de examencommissie, het uitvoeren van een promotieonderzoek naast de geneeskunde-opleiding is alleen mogelijk in een kader waarbij de opleiding een constante, gewaarborgde factor is naast de grilligheid van wetenschappelijk onderzoek. Ik wil u allen, en in het bijzonder Prof. dr. Laan, bedanken voor het bieden van dit kader en de kans om dit traject te doorlopen.

Sander en Martijn, ik vind het een eer dat ik met jullie aan mijn zijde dit proefschrift mag verdedigen. Sander, samen begonnen aan de studie biologie en na verschillende omwegen zullen we in de toekomst ons toch beiden weer gaan bezighouden met het bewegingsapparaat, ik kijk uit naar de vele discussies met een Belgisch biertje die daarop gaan volgen. Martijn, het afgelopen jaar was ik niet meer de enige Broen waarvan een referentie verscheen op Pubmed, daar was jij mede verantwoordelijk voor. Ik vind het fijn dat je me bijstaat als paranimf en stel voor de eerste neurologische-reumatologische overlap ziekte als gezamenlijk onderzoeksproject aan te grijpen!

Beste ouders, zus, schoonfamilie, oma en opa, bedankt voor de onvoorwaardelijke steun, de interesse en het vertrouwen dat ik van jullie gekregen heb tijdens mijn studie en promotieonderzoek. Het is de kern geweest van het vertrouwen dat ik nodig had om beide te volbrengen.

Lieve Christel, bedankt voor je onmisbare steun, begrip en geduld als ik weer eens een avond of weekend meer aandacht had voor SNPs dan voor jou, of een periode in het buitenland was. Het lijkt me toepasselijk om dit proefschrift met een ietwat aangepaste citatie van Oscar Wilde te eindigen: *“Love is more important than science, because it needs no explanation”.*

Curriculum Vitae

De auteur van dit proefschrift, Jacobus (Jasper) Christiaan Andreas Broen werd in de paasnacht, 7 april 1985, geboren te Linne (in Limburg). In Linne heeft hij het grootste deel van zijn jeugd doorgebracht, waarna hij in 2003 zijn VWO diploma behaalde aan het Stedelijk Lyceum te Roermond. Doordat hij door de *numerus fixus* niet meteen kon beginnen met de studie Geneeskunde, heeft hij eerst een jaar Medische Biologie gestudeerd aan de Radboud Universiteit te Nijmegen. Nadat in het eerste jaar de propedeuse behaald werd, kon hij in 2004 beginnen met de studie Geneeskunde, eveneens te Nijmegen. Tijdens de eerste jaren van de studie was hij betrokken als student-assistent bij onderwijs van de afdeling Pathologie. In 2007 startte hij als student-assistent bij de afdeling Reumatische Ziekten in het UMC St. Radboud, alwaar hij onder leiding van Dr. Vonk en Dr. Radstake zich bezig houdt met het, voor wetenschappelijk onderzoek, verzamelen van bloed afkomstig van Systemische Sclerose patiënten. Geleidelijk raakt hij betrokken bij een samenwerking tussen Dr. Coenen (afdeling Antropogenetica) en Dr. Radstake, betreffende het opbouwen van een genetische en klinische database van systemische sclerose patiënten. Bij dit project, dat binnen korte tijd is uitgegroeid tot een wereldwijde samenwerking tussen reumatologen, is hij tot op heden betrokken en het vormt de basis voor dit proefschrift. In 2008 heeft hij onder leiding van Dr. Radstake en Prof. Lafyatis een half jaar een wetenschappelijke stage doorlopen op de afdeling Reumatologie in het Boston University Medical Center, te Boston, USA. Na deze stage werd in overleg met de examencommisie besloten om het wetenschappelijk onderzoek voort te zetten als promotietraject parallel aan de laatste klinische jaren (coschappen) van de opleiding Geneeskunde. Gedurende zijn promotietraject heeft hij diverse internationale presentaties gegeven en awards in ontvangst mogen nemen. De laatste maanden van zijn promotietraject heeft hij als visiting academic researcher doorgebracht aan The University of Glasgow (Glasgow, Groot-Brittannie) met als behaald doel het opstarten van een samenwerking aangaande bio-ageing research in reumatische ziekten. In 2013 begint hij aan de specialisatie reumatologie.

Curriculum Vitae (English)

The author of this thesis, Jacobus (Jasper) Christiaan Andreas Broen was born on Easter Sunday, April 7th, 1985, in Linne (Limburg, Netherlands). He spent most of his childhood in Linne, after which he graduated from high school in 2003 at the Stedelijk Lyceum in Roermond. He could not immediately start studying medicine after high school because of the limited enrollment and he spent a year studying Medical Biology at the Radboud University Nijmegen. After this year he was admitted to start with studying Medicine, also in Nijmegen. During the first years of the study he was involved as a student assistant in the Pathology Department of Education. In 2007 he started as a student assistant at the Department of Rheumatic Diseases in the UMC St. Radboud where he collected blood from systemic sclerosis patients for scientific research, under supervision of Dr. Vonk and Dr. Radstake. He gradually became involved in a collaboration between Dr. Coenen (Department of Human Genetics) and Dr. Radstake, regarding the construction of a genetic and clinical database of systemic sclerosis patients. This project evolved into a global collaboration between rheumatologists and it forms the basis for this thesis. In 2008 he underwent a scientific traineeship of six months at the Department of Rheumatology at the Boston University Medical Center, in Boston, USA supervised by Dr. Radstake and Prof. Lafyatis. After this traineeship, Dr. Radstake and he decided in consultation with the exam committee of the medical faculty in Nijmegen, to start working as a PhD student in parallel with the clinical internships in the final three years of medical training. During his time as a PhD student he had the opportunity to present his work at several international congresses and had the honor to receive scientific awards for the research conducted. He finished his thesis as a visiting academic researcher at The University of Glasgow (Glasgow, Great Britain) where he focused on bio-aging research in rheumatic diseases. In 2013 he will start a six year training-program to become a rheumatologist.

Publication list

1. Rueda B, **Broen J**, Torres O, Simeon C, Ortego-Centeno N, Schrijvenaars MM, Vonk MC, Fonollosa V, van den Hoogen FH, Coenen MJ, Sanchez-Román J, Aguirre-Zamorano MA, García-Portales R, Pros A, Camps MT, Gonzalez-Gay MA, Martin J, Radstake TR. The interleukin 23 receptor gene does not confer risk to systemic sclerosis and is not associated with systemic sclerosis disease phenotype. *Ann Rheum Dis. 2009 Feb;68(2): 253-6.*

2. Rueda B, **Broen J**, Simeon C, Hesselstrand R, Diaz B, Suárez H, Ortego-Centeno N, Riemekasten G, Fonollosa V, Vonk MC, van den Hoogen FH, Sanchez-Román J, Aguirre-Zamorano MA, García-Portales R, Pros A, Camps MT, Gonzalez-Gay MA, Coenen MJ, Airo P, Beretta L, Scorza R, van Laar J, Gonzalez-Escribano MF, Nelson JL, Radstake TR, Martin J. The STAT4 gene influences the genetic predisposition to systemic sclerosis phenotype. *Hum Mol Genet. 2009 Jun 1;18(11):2071-7.*

3. Radstake TR, van Bon L, **Broen J**, Hussiani A, Hesselstrand R, Wuttge DM, Deng Y, Simms R, Lubberts E, Lafyatis R. The pronounced Th17 profile in systemic sclerosis (SSc) together with intracellular expression of TGFbeta and IFNgamma distinguishes SSc phenotypes. *PLoS One. 2009 Jun 17;4(6):e5903.*

4. Radstake TR, van Bon L, **Broen J**, Wenink M, Santegoets K, Deng Y, Hussaini A, Simms R, Cruikshank WW, Lafyatis R. Increased frequency and compromised function of T regulatory cells in systemic sclerosis (SSc) is related to a diminished CD69 and TGFbeta expression. *PLoS One. 2009 Jun 22;4(6):e5981.*

5. Rueda B, Gourh P, **Broen J**, Agarwal SK, Simeon C, Ortego-Centeno N, Vonk MC, Coenen M, Riemekasten G, Hunzelmann N, Hesselstrand R, Tan FK, Reveille JD, Assassi S, Garcia-Hernandez FJ, Carreira P, Camps M, Fernandez-Nebro A, Garcia de la Peña P, Nearney T, Hilda D, Gónzalez-Gay MA, Airo P, Beretta L, Scorza R, Radstake TR, Mayes MD, Arnett FC, Martin J. BANK1 functional variants are associated with susceptibility to diffuse systemic sclerosis in Caucasians. *Ann Rheum Dis. 2010 Apr;69(4):700-5. Epub 2009 Oct 8.*

6. Wenink MH, Santegoets KC, **Broen JC**, van Bon L, Abdollahi-Roodsaz S, Popa C, Huijbens R, Remijn T, Lubberts E, van Riel PL, van den Berg WB, Radstake TR. TLR2 promotes Th2/Th17 responses via TLR4 and TLR7/8 by abrogating the type I IFN amplification loop. *J Immunol. 2009 Dec 1;183(11): 6960-70. Epub 2009 Nov 13.*

7. **Broen J**, Gourh P, Rueda B, Coenen M, Mayes M, Martin J, Arnett FC, Radstake TR*;* European Consortium on Systemic Sclerosis Genetics. The FAS -670A>G polymorphism influences susceptibility to systemic sclerosis phenotypes. *Arthritis Rheum. 2009 Dec;60(12):3815-20.*

8. Radstake TR, Gorlova O, Rueda B, Martin JE, Alizadeh BZ, Palomino-Morales R, Coenen MJ, Vonk MC, Voskuyl AE, Schuerwegh AJ, **Broen JC**, van Riel PL, van 't Slot R, Italiaander A, Ophoff RA, Riemekasten G, Hunzelmann N, Simeon CP, Ortego-Centeno N, González-Gay MA, González-Escribano MF; Spanish Scleroderma Group, Airo P, van Laar J, Herrick A, Worthington J, Hesselstrand R, Smith V, de Keyser F, Houssiau F, Chee MM, Madhok R, Shiels P, Westhovens R, Kreuter A, Kiener H, de Baere E, Witte T, Padykov L, Klareskog L, Beretta L, Scorza R, Lie BA, Hoffmann-Vold AM, Carreira P, Varga J, Hinchcliff M, Gregersen PK, Lee AT, Ying J, Han Y, Weng SF, Amos CI, Wigley FM, Hummers L, Nelson JL, Agarwal SK, Assassi S, Gourh P, Tan FK, Koeleman BP, Arnett FC, Martin J, Mayes MD. Genome-wide association study of systemic sclerosis identifies CD247 as a new susceptibility locus. *Nat Genet. 2010 May;42(5):426-9.*

9. Alizadeh BZ*, **Broen J***, Rueda B, Hesselstrand R, Wuttge D, Simeon C, Ortego-Centeno N, Gonzalez-Gay MA, Pros A, Herrick A, Worthington J, Denton C, Fonseca C, Riemekasten G, Vonk MC, van den Hoogen F, Guiducci S, Matucci-Cerinic M, Scorza R, Beretta L, Airó P, Coenen M, Martin J, Koeleman BP, Radstake TR; EUSTAR. Functional variants of Fc gamma receptor (FCGR2A) and FCGR3A are not associated with susceptibility to systemic sclerosis in a large European Study (EUSTAR). *J Rheumatol. 2010 Aug 1;37(8):1673-9.*

10. **Broen JC**, Wolvers-Tettero IL, Geurts-van Bon L, Vonk MC, Coenen MJ, Lafyatis R, Radstake TR, Langerak AW. Skewed X chromosomal inactivation impacts T regulatory cell function in systemic sclerosis. *Ann Rheum Dis. 2010 Aug 10.*

11. Diaz-Gallo L, Gourh P, **Broen J**, Simeon C, Fonollosa V, Ortego-Centeno N, Agarwal S, Vonk M, Coenen M, Riemekasten G, Hunzelmann N, Hesselstrand R, Tan F, Reveille J, Assassi S, García-Hernandez F, Carreira P, Camps M, Fernandez-Nebro A, de la Peña PG, Nearney T, Hilda D, González-Gay M, Airo P, Beretta L, Scorza R, Herrick A, Worthington J, Pros A, Gómez-Gracia I, Trapiella L, Espinosa G, Castellvi I, Witte T, de Keyser F, Vanthuyne M, Mayes M, Radstake T, Arnett F, Martin J, Rueda B. Analysis of the influence of PTPN22 gene polymorphisms in systemic sclerosis. *Ann Rheum Dis. 2010 Dec 3.*

12. **Broen J**, Coenen MJH, Rueda B,Witte T, Padyukov L, Klareskog L, Hesselstrand R, Wuttge DM, Simeon CP, Ortego-Centeno N, Gonzalez-Gay M, Pros A, Hunzelman N, Riemekasten G, Kreuter A, Vonk M, Scorza R, Beretta L, Airo' P, Van Riel PLCM, Kimberly R, Martin J, Edberg J, Radstake TRDJ. The functional polymorphism in FcαRI (CD89) does not contribute to Systemic Sclerosis or Rheumatoid Arthritis susceptibility. *J Rheumatol. 2010 Dec 15.*

13. Bossini-Castillo L, **Broen J**, Simeon CP, Beretta L, Vonk MC, Ortego-Centeno N, Espinosa G, Carreira P, Camps MT, Navarrete N, González-Escribano MF, Vicente-Rabaneda E, Rodríguez L, Tolosa C, Román-Ivorra JA, Gómez-Gracia I, García-Hernández FJ, Castellví I, Gallego M, Fernández-Nebro A, Egurbide MV, Follonosa V, De la Peña PG, Pros A,

González-Gay MA, Hesselstrand R, Riemekasten G, Witte T, Coenen MJH, Koeleman BPC, Houssiau F, Smith V, De Keyser F, Westhovens R, Lories R, Voskuyl AE, Scheurwegh AJ, Chee MM, Madhok R, Shiels P, Kiener H, De Baere E, Padykov L, Klareskog L, Hoffman-Vold AM, Lie BA, Airó P, Scorza R, Van Laar, Hunzelmann N, Kreuter A, Herrick A, Worthington J, Radstake TRDJ, Martin J*, Rueda B*. A replication study confirms the association of TNFSF4 (OX40L) polymorphisms with Systemic Sclerosis in a large European cohort. *Ann Rheum Dis. 2010 Dec 27.*

14. **Broen JCA**, Radstake TRDJ. How birds of a feather flock together: Genetics in autoimmune diseases. *Expert Rev Clin Immunol. 2011 Mar;7(2):127-8.*

15. Gorlova O, Martin JE, Rueda B, Koeleman BPC, Ying J, Teruel M, Diaz-Gallo LM, **Broen J**, Vonk MC, Simeon CP, Alizadeh BZ, Coenen MJH, Voskuyl AE, Schuerwegh AJ, Van Riel PLCM, Vanthuyne M, Van 't Slot R, Italiaander, Ophoff RA, Hunzelmann N, Fonollosa V, Ortego-Centeno N, González-Gay MA, García-Hernández FJ, González-Escribano MF, Airo' P, Van Laar J, Worthington J, Hesselstrand R, Smith V, De Keyser F, Houssiau F, Chee MM, Madhok R, Shiels P, Westhovens R, Kreuter A, De Baere E, Witte T, Padyukov L, Nordin A, Scorza R, Lunardi C, Lie BA, Hoffmann-Vold AM, De la Peña PG, Carreira P, Varga J, Hinchcliff M, Lee AT, Gourh P, Amos CI, Wigley FM, Hummers LK, Nelson JL, Riemekasten G, Herrick A, Beretta L, Fonseca C, Denton CP, Gregersen PK, Agarwal S, Assassi S, Tan FK, Arnett FC, Radstake TRDJ, Mayes MD, Martin J. Identification of novel genetic markers associated with clinical phenotypes and autoantibody subsets of systemic sclerosis through a genome wide association strategy. *PLoS Genet. 2011 Jul;7(7):e1002178. Epub 2011 Jul 14*

16. **Broen JCA**, Radstake TRDJ, Coenen MJH. Deciphering the genetic background of systemic sclerosis. *Expert Rev Clin Immunol. 2011 Jul;7(4): 449-62.*

17. **Broen JCA**, Bossini-Castillo L, Van Bon L, Vonk MC, Knaapen H, Beretta L, Rueda B, Hesselstrand R, Herrick A, Worthington J, Hunzelman N, Denton C, Fonseca C, Riemekasten G, Kiener H, Scorza R, Simeon CP, Ortego-Centeno N (for the Spanish Systemic Sclerosis group), Gonzalez-Gay MA, Airo' P, Coenen MJH, Martin J, and Radstake TRDJ. A rare polymorphism in Toll Like Receptor 2 is associated with systemic sclerosis phenotype and increases production of inflammatory mediators. *Arthritis Rheum. 2011 Sep 8. doi: 10.1002/art.33325.*

18. **Broen JCA**, Dieude P, Vonk MC, Beretta L, Carmona FD, Herrick A, Worthington J, Hunzelmann N, Riemekasten G, Kiener H, Scorza R, Simeon CP, Fonollosa V (for the Spanish Systemic Sclerosis group), Carreira P, Ortego-Centeno N, Gonzalez-Gay MA, Airo' P, Coenen MJH, Tsang, K, Aliprantis AO, Martin J,* Allanore Y*, Radstake TRDJ*. Polymorphisms in the Interleukin 4, Interleukin 13 and corresponding receptor genes are not associated with Systemic Sclerosis and do not influence gene expression. *J Rheumatol. 2011 Nov 1.*

19. Bossini-Castillo L, Martin JE, **Broen JCA**, Simeon CP, Beretta L, Vonk MC, Callejas JL, Castellví I, Carreira P, García-Hernández FJ, Fernandez de Castro M, and the Spanish Scleroderma Group, Coenen MJH, Riemekasten G, Witte T, Hunzelman N, Kreuter A, Distler JWH, Koeleman BP, Voskuyl AE, Schuerwegh AJ, Palm Ø, Hesselstrand R, Nordin A, Airó P, Lunardi C, Scorza R, Shiels PG, van Laar JM, Herrick A, Worthington J, Denton C, Tan FK, Arnett FC, Agarwal SK, Assassi S, Fonseca C, Mayes MD, Radstake TRDJ, Martin J. A GWAS follow-up study reveals the association of IL12RB2 gene with Systemic Sclerosis in Caucasian populations. *Hum Mol Genet. 2011 Nov 29.*

20. **Broen JCA**, Coenen MJH, Radstake TRDJ. Genetics in Systemic Sclerosis: an update. *Curr Rheumatol Rep. 2011 Nov 19*

21. Martin JE*, Carmona F*, **Broen JCA**, Simeon CP, Vonk MC, Carreira P, Ríos-Fernández R, Espinosa G, Vicente-Rabaneda E, Tolosa C, García-Hernández FJ, Castellví I, Fonollosa F, González-Gay MA, Sáez-Comet L, Portales RG, Garcia de la Peña G, Andreu JL, Diaz B, Martinez-Estupiñan L, and Spanish Scleroderma Group, Voskuyl AE, Schuerwegh AJ, Vanthuyne A, Smith V, De Langhe E, Riemekasten G, Witte T, Hunzelmann N, Kreuter A, Palm Ø, Chee MM, van Laar JM, Denton C, Herrick A, Worthington J, Koeleman BPC, Radstake TRDJ*, Fonseca C*, Martín J*. The autoimmune disease-associated IL2RA locus is involved in the clinical manifestations of systemic sclerosis. *Genes Immun. 2011 Oct 20*

22. **Broen JCA**, McGlynn L, Shiels PG, Radstake TRDJ. Bioageing research in systemic sclerosis: time to grow up? (Book Chapter, not peer-reviewed) *Systemic Sclerosis - An Update on the Aberrant Immune System and Clinical Features, ISBN: 978-953-307-869-4*

23. Martin JE, **Broen JC**, Carmona FD, Teruel M, Simeon CP, Vonk MC, van 't Slot R, Rodriguez-Rodriguez L, Vicente E, Fonollosa V, Ortego-Centeno N, González-Gay MA, García-Hernández FJ, de la Peña PG, Carreira P; Spanish Scleroderma Group, Voskuyl AE, Schuerwegh AJ, van Riel PL, Kreuter A, Witte T, Riemekasten G, Airo P, Scorza R, Lunardi C, Hunzelmann N, Distler JH, Beretta L, van Laar J, Chee MM, Worthington J, Herrick A, Denton C, Tan FK, Arnett FC, Assassi S, Fonseca C, Mayes MD, Radstake TR, Koeleman BP, Martin J. Identification of CSK as a systemic sclerosis genetic risk factor through Genome Wide Association Study follow-up*. Hum Mol Genet. 2012 Mar 22.*

24. Bossini-Castillo L, Simeon CP, Beretta L, **Broen JCA**, Vonk MC, Ríos-Fernández R, Espinosa G, Carreira P, Camps MT, M, Castillo J, González-Gay MA, Beltrán E, del Carmen Freire M, Narváez FJ , Tolosa C, the Spanish Scleroderma Group #, Witte T, Kreuter A, Schuerwegh AJ, Hoffmann-Vold AM, Hesselstrand R, Lunardi C, van Laar JM, Chee MM, Herrick A, Koeleman BP, Fonseca C, Radstake TRDJ*, Martin J* A multicentre study confirms CD226 gene association with Systemic Sclerosis-related pulmonary fibrosis. *Arthritis Res Ther. 2012 Apr 24;14(2):R85*

25. McKinney C, **Broen JCA**, Vonk MC, Beretta L, Hesselstrand R, Hunzelman N, Riemekasten G, Scorza R, Simeon CP, Fonollosa V, Carreira P, Ortego-Centeno N, Gonzalez-Gay MA, Airo' P, Martin JE*, Radstake TRDJ*, Merriman TR*. Evidence that deletion at FCGR3B is a risk factor for systemic sclerosis. *Genes Immun. 2012 May 3. doi: 10.1038*

26. Teruel M, Simeon CP, **Broen** J, Vonk MC, Carreira P, Camps MT, García-Portales R, Delgado-Frías E, Gallego M, Espinosa G; the Spanish Scleroderma Group, Beretta L, Airó P, Lunardi C, Riemekasten G, Witte T, Krieg T, Kreuter A, Distler JH, Hunzelmann N, Koeleman BP, Voskuyl AE, Schuerwegh AJ, González-Gay MA, Radstake TR, Martin J. Analysis of the association between CD40 and CD40 ligand polymorphisms and systemic sclerosis. *Arthritis Res Ther. 2012 Jun 25;14(3):R154. [Epub ahead of print]*

27. Bossini-Castillo L, Martin JE, **Broen J**, Simeon CP, Beretta L, Gorlova OY, Vonk MC, Ortego-Centeno N, Espinosa G, Carreira P, García de la Peña P, Oreiro N, Román-Ivorra JA, Castillo MJ, González-Gay MA, Sáez-Comet L, Castellví I, Schuerwegh AJ, Voskuyl AE, Hoffmann-Vold AM, Hesselstrand R, Nordin A, Lunardi C, Scorza R, van Laar JM, Shiels PG, Herrick A, Worthington J, Fonseca C, Denton C, Tan FK, Arnett FC, Assassi S, Koeleman BP, Mayes MD, Radstake TR, Martin J; the Spanish Scleroderma Group. Confirmation of TNIP1 but not RHOB and PSORS1C1 as systemic sclerosis risk factors in a large independent replication study. *Ann Rheum Dis. 2012 Aug 15. [Epub ahead of print]*

National publications (not peer-reviewed)

1. Radstake TRDJ & Broen JCA Trans-Atlantisch onderzoek leidt tot identificatie van genen als veroorzaker van Systemische Sclerose. *Nederlands Tijdschrift voor Reumatologie 2010.*

2. Broen JCA & Radstake TRDJ Welke telomeerziekten kent de reumatoloog? *Spreekuur Reumatologie. Bohn Stafleu en van Loghum 2012.*

NB: *Deze personen delen de auteurspositie op basis van een gelijkwaardige bijdrage aan de publicatie.

Appendix I: Collaborations

This thesis would not have been possible to create without the help from many rheumatologists and researchers, and most important the patients and healthy controls that participated. Below are all the persons that I would like to thank for their participation. I had the pleasure to meet the majority of them during my recent years as a PhD student and some of them even became friends. Hopefully this list will keep on growing to tackle the difficult obstacles in SSc basic research in a strong, combined effort.

Name	Affilliation
Aguirre-Zamorano MA	Hospital Reina Sofía, Córdoba, Spain
Airo' P	Servizio di Reumatologia ed Immunologia Clinica, Spedali Civili, Brescia, Italia
Aliprantis AO	Department of Immunology and Infectious Diseases, Harvard School of Public Health, Boston, MA 02115, USA
Allanore Y	Université Paris Descartes, INSERM U1016, Hôpital Cochin, Service de Rhumatologie A, AP-HP, Paris, France
Arnett FC	Division of Rheumatology and Clinical Immunogenetics, Department of Internal Medicine, University of Texas Health Science Center at Houston (UTHSC-H), Houston, TX, USA
Beretta L	Referral Center for Systemic Autoimmune Diseases, University of Milan, Italy.
Bossini-Castillo L	Instituto de Parasitología y Biomedicina, CSIC, Granada, Spain
Broen JCA	Dept of Rheumatology, Radboud University Nijmegen Medical Center, The Netherlands
Brouwer C	Dept of Rheumatology, Radboud University Nijmegen Medical Center, The Netherlands
Camps MT	Hospital Carlos Haya, Málaga, Spain
Carmona FD	Instituto de Parasitología y Biomedicina, CSIC, Granada, Spain
Carreira P	Servicio de Reumatología, Hospital 12 de Octubre, Madrid, Spain
Castellví I	Department of Rheumatology, Hospital Sant Pau, Barcelona
Chee MM	University of Glasgow, Glasgow, United Kingdom

Name	Affilliation
Claes K	Department of Genetics, University of Ghent, Ghent, Belgium
Coenen MJH	Department of Human Genetics, Radboud University Nijmegen Medical Center, The Netherlands.
Cossu M	Dept of Rheumatology, Radboud University Nijmegen Medical Center, The Netherlands.
De Keyser F	University of Ghent, Ghent, Belgium
De la Peña PG	Servicio de Reumatología, Hospital Ramón y Cajal, Madrid, Spain
De Langhe E	University of Leuven (KULeuven), Leuven, Belgium
Denton C	Centre for Rheumatology, Royal Free and University College Medical School, London, United Kingdom
Diaz B	Servicio de Medicina Interna, Hospital Universitario Central de Asturias, Oviedo, Spain
Dieude P	Universite´ Diderot Paris 7, Service de Rhumatologie, Hospital Bichat Claude Bernard, Paris, France
Edberg J	Division of Clinical Immunology and Rheumatology, Department of Medicine, University of Alabama at Birmingham, Birmingham, AL, USA
Egurbide MV	Servicio de Medicina Interna, Hospital de Cruces, Barakaldo, Spain
Espinosa G	Servicio de Medicina Interna, Hospital Clínico de Barcelona, Barcelona, Spain
Fernández-Nebro A	Servicio de Reumatología, Hospital Carlos Haya, Málaga, Spain
Fonollosa V	Servicio de Medicina Interna, Hospital Valle de Hebron, Barcelona, Spain
Fonseca C	Centre for Rheumatology, Royal Free and University College Medical School, London, United Kingdom
Gallego M	Servicio de Medicina Interna, Hospital Central de Asturias, Oviedo, Spain
García-Hernández FJ	Servicio de Medicina Interna, Hospital Virgen del Rocío, Sevilla, Spain
García-Portales R	Servicio de Reumatología, Hospital Virgen de la Victoria, Málaga, Spain
Gómez-Gracia I	Servicio de Reumatología, Hospital Reina Sofía, Córdoba, Spain

Name	Affilliation
González-Escribano MF	Servicio de Inmunología, Hospital Virgen del Rocío, Sevilla, Spain
Gonzalez-Gay MA	Servicio de Reumatologia, Hospital Marques de Valdecillas, Santander, Spain.
Gourh P	Division of Rheumatology and Clinical Immunogenetics, Department of Internal Medicine, University of Texas Health Science Center at Houston (UTHSC-H), Houston, TX, USA
Herrick A	Rheumatic Diseases Centre, University of Manchester, Salford Royal NHS Foundation Trust, UK
Hesselstrand R	Department of Rheumatology, Lund University Hospital, Lund, Sweden
Houssiau F	University of Leuven, Belgium
Hunzelmann N	Department of Dermatology, University of Cologne, Germany
Kiener H	Department of Internal Medicine, division of Rheumatology, University of Vienna, Austria
Kimberly R	Division of Clinical Immunology and Rheumatology, Department of Medicine, University of Alabama at Birmingham, Birmingham, AL, USA
Klareskog L	Karolinska Institute, Stockholm, Sweden
Knaapen H	Dept of Rheumatology, Radboud University Nijmegen Medical Center, The Netherlands
Koeleman BP	Section Complex Genetics, Department of Medical Genetics, University Medical Center Utrecht, Utrecht, The Netherlands
Kreuter A	Ruhr University of Bochum, Bochum, Germany
Kyburz D	University Hospital Zurich Div. of Rheumatology Zurich, Switzerland
Lafyatis R	Boston University Medical Center Boston, USA
Langerak AW	Department of Immunology, Erasmus MC, University Medical Center, Rotterdam, Netherlands
Lie BA	Institute of Immunology, Rikshospitalet, Oslo University Hospital, Oslo, Norway
Lories R	University of Antwerpen, Antwerpen, Belgium
Madhok R	University of Glasgow, Glasgow, United Kingdom

Name	Affilliation
Martin J	Instituto de Parasitología y Biomedicina, CSIC, Granada, Spain
Mayes M	Division of Rheumatology and Clinical Immunogenetics, Department of Internal Medicine, University of Texas Health Science Center at Houston (UTHSC-H), Houston, TX, USA
Mayo M	Department of Computer Science, Waikato University, Hamilton, New Zealand
McGlynn L	University of Glasgow, Glasgow, United Kingdom
Navarrete N	Servicio de Medicina Interna, Hospital Virgen de las Nieves, Granada, Spain
Nelson JL	Department of Medicine, University of Washington, Seattle, WA, USA
Niederer F	University Hospital Zurich Div. of Rheumatology Zurich, Switzerland
Nordin A	Karolinska Institute, Stockholm, Sweden
Ortego-Centeno N	Servicio de Medicina Interna, Hospital Clínico Universitario, Granada, Spain
Padyukov L	Karolinska Institute, Stockholm, Sweden
Palm O	Department of Rheumatology, Rikshospitalet, Oslo University Hospital, Oslo, Norway
Pros A	Servicio de Reumatología, Hospital Del Mar, Barcelona, Spain
Radstake TRDJ	Dept of Rheumatology, Radboud University Nijmegen Medical Center, The Netherlands
Riemekasten G	Dept of Rheumatology and Clinical Immunology, Charité University Hospital, Berlin and German Rheumatism Research Centre, a Leibniz institute, Germany
Rodríguez L	Servicio de Reumatología, Hospital Clinico San Carlos, Madrid, Spain
Román-Ivorra JA	Servicio de Reumatología, Hospital del Doctor Peset aleixandre, Valencia, Spain
Rueda B	Instituto de Parasitología y Biomedicina, CSIC, Granada, Spain
Sanchez-Román J	Servicio de Medicina Interna, Hospital Virgen del Rocio, Sevilla, Spain
Schuerwegh AJ	Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands

Name	Affilliation
Scorza R	Referral Center for Systemic Autoimmune Diseases, University of Milan, Italy
Shiels P	University of Glasgow, Glasgow, United Kingdom
Simeon CP	Servicio de Medicina Interna, Hospital Valle de Hebron, Barcelona, Spain
Smith V	University of Ghent, Ghent, Belgium. 31 University of Antwerpen, Antwerpen, Belgium
Suárez H	Servicio de Medicina Interna, Hospital Universitario Central de Asturias, Oviedo, Spain
Tolosa C	Servicio de Medicina Interna, Hospital Parc Tauli, Sabadell, Spain
Tsang K	Department of Immunology and Infectious Diseases, Harvard School of Public Health, Boston, MA 02115, USA
van Bon L	Dept of Rheumatology, Radboud University Nijmegen Medical Center, The Netherlands
van den Hoogen FH	Dept of Rheumatology, St Maartenskliniek,, The Netherlands
van Laar JM	Institute of Cellular Medicine, Newcastle University, Newcastle, UK
van Riel PL	Dept of Rheumatology, Radboud University Nijmegen Medical Center, The Netherlands.
Vicente-Rabaneda E	Department of Rheumatology, Hospital La Princesa, Madrid
Vonk MC	Dept of Rheumatology, Radboud University Nijmegen Medical Center, The Netherlands.
Voskuyl AE	Department of Rheumatology, VU University Medical Center, Amsterdam, The Netherlands
Westhovens R	University of Antwerpen, Antwerpen, Belgium
Witte T	Department of Medicine, Clinic for Immunology and Rheumatology, Hannover Medical School, Hannover, Germany
Wolvers-Tettero ILM	Department of Immunology, Erasmus MC, University Medical Center, Rotterdam, Netherlands
Worthington J	Rheumatic Diseases Centre, University of Manchester, Salford Royal NHS Foundation Trust, UK
Wuttge DM	Department of Medicine, Division of Rheumatology at Karolinska University Hospital, Stockholm, Sweden

Appendix II: Abbrevations

A	Adenine
ACA	Anti Centromere Antibodies
ACR	American College of Rheumatologists
AID	Auto-immune disease(s)
ANA	Anti Nuclear Antibodies
anti-CCP	Anti Cyclic Citrullinated Peptide
anti-scl70	Anti-topoisomerase antibodies
anti-topo	Anti-topoisomerase antibodies
APC	Antigen Presenting Cell(s)
ATA	Anti-topoisomerase antibodies
BANK1	B-Cell Scaffold Protein with Ankyrin repeats
BC	Before Christ
BCR	B Cell Receptor
BD	Breslow-Day
BDCA	Blood dendritic cell antigen
BLK	B lymphocyte specific tyrosine kinase
C	Cytosine
CA	California
CD	Cluster of differentiation
CDH7	Cadherin 7
CI	Confidence Interval
cM	centiMorgan
CTGF	Connective Tissue Growth Factor
CTLA4	Cytotoxic T Lymphocyte-associated 4
dcSSc	diffuse cutaneous Systemic Sclerosis
DLCO	diffusion capacity of the lung for carbon monoxide
DNA	Deoxyribonucleic acid
DZ	Dyzigotic
ELISA	Enzyme-linked immunosorbent assay
EXOC2	Exocyst complex component 2
FAS	TNF receptor superfamily, member 6
FBN	Fibrillin 1
FcαR1	receptor for Fc fragment of IgA
FOXP3	Forkhead Box P3
FVC	Forced vital capacity

G	Guanine
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GRB10	Growth Factor Receptor-Bound protein 10
GWAS	Genome Wide Association Study
HAPMAP	haplotype map
HLA	Human Leukocyte Antigen
HUMARA	Human Androgen Receptor
HWE	Hardy-Weinberg Equilibrium
ICAM-1	intercellular adhesion molecule 1
IFN	Interferon
Ig	Immunoglobulin
IL	Interleukin
IL13RA1	Interleukin 13 receptor alpha 1
IL4R	Interleukin 4 receptor
IRAK1	Interleukin 1 Receptor-Associated Kinase 1
IRF	Interferon Regulatory Factor
JAK	Janus Kinase
lcSSc	limited cutaneous Systemic Sclerosis
LD	linkage disequilibrium
M-H	Mantel-Haenzel
MAF	minor allele frequency
MDR	Multi Dimensional Reduction
MHC	Major Histocompatibiliy Complex
moDC	monocyte derived dendritic cells
mRNA	messenger Ribonucleic Acid
MyDC	myeloid dendritic cells
MZ	Monozygotic
n	Number
NLRP1	NLR Family Pyrin Domain Containing 1
NOTCH	Notch homolog 1, translocation associated
OR	Odds Ratio
OX40L	Tumor Necrosis Factor Receptor Superfamily, member 4
PAH	Pulmonary Arterial Hypertension
PBEF	Pre-B-Cell Colony-Enhancing Factor 1
PBL	Peripheral Blood Leukocytes
PBMC	Peripheral Blood Mononuclear Cells
Pcor	Corrected p value

PCR	Polymerase Chain Reaction
PDC	Plasmacytoid Dendritic Cells
PDGF	Platelet Derived Growth Factor
POL	RNA polymerase
PTPN22	Protein Tyrosine Phosphatase Nonreceptor-type 22
Q-PCR	Quantitative PCR
RA	Rheumatoid Arthritis
RE	Relative Expression
RF	Rheumatoid factor
RFLP	restriction fragment length polymorphism
RNA	Ribonucleic Acid
RR	Relative Risk
SD	Standard Deviation
SLE	Systemic Lupus Erythematosus
SNP	Single Nucleotide Polymorphism
SOX5	SRY-BOX 5
SPARC	Secreted Protein Acidic Cystine-Rich
SSc	Systemic Sclerosis
STAT4	Signal Transducer and Activator of Transcription 4
T	Thymine
T/S	telomere repeat copy number to single copy gene number
TBX21	T-Box 21
TGF	Transforming Growth Factor
Th	T helper
TLR	Toll Like Receptor
TNF	Tumor Necrosis Factor
TNFSF4	Tumor Necrosis Factor Ligand Superfamily Member 4
TNIP	TNFAIP3 interacting protein 1
TNPO3	Transportin 3
TOPO1	Topoisomerase
Tregs	T reulatory cells
TYK	Tyrosyne Kinase
UPAR	UPA Receptor
USA	United States of America
UTR	Untranslated Region
XCI	X Chromosomal Inactivation

Publication of this thesis was made possible by: